

Mon Apr 7 07:36:48 2003

us-10-057-951-2.rpr

age 1

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 06:47:33 ; Search time 17.6747 Seconds

(without alignments)
1430.480 Million cell updates/sec

Title: US-10-057-951-2

Sequence: 1 MLAWQAFVLSNMMLAEV.....PVDQSGSTPLMGQACTGCA 263

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 134.5 | 11.1 | 562 | 1 UKHUT | t-plasminogen acti |
| 2 | 134.5 | 11.1 | 359 | 1 A35029 | t-plasminogen acti |
| 3 | 147.2 | 10.6 | 251 | 1 A35029 | t-plasminogen acti |
| 4 | 146.2 | 10.6 | 251 | 1 A35029 | t-plasminogen acti |
| 5 | 145.2 | 10.2 | 231 | 1 B3098 | hepatocyte growth |
| 6 | 142 | 10.2 | 810 | 1 P1HU | plasmin (EC 3.4.21 |
| 7 | 138 | 9.9 | 169 | 2 A40522 | plasmin (EC 3.4.21 |
| 8 | 138 | 9.9 | 433 | 1 UKBAY | u-plasminogen acti |
| 9 | 137 | 9.8 | 716 | 1 JC5061 | macrophage-stimula |
| 10 | 135.5 | 9.7 | 442 | 1 UKPG | u-plasminogen acti |
| 11 | 134 | 9.6 | 431 | 2 JS0599 | u-plasminogen acti |
| 12 | 134 | 9.6 | 477 | 1 A34369 | t-plasminogen acti |
| 13 | 134 | 9.6 | 477 | 2 JS0598 | t-plasminogen acti |
| 14 | 134 | 9.6 | 716 | 1 A40332 | macrophage-stimula |
| 15 | 132 | 9.5 | 394 | 2 JS0600 | u-plasminogen acti |
| 16 | 132 | 9.5 | 431 | 1 UKHUT | u-plasminogen acti |
| 17 | 130 | 9.3 | 810 | 2 A46260 | plasmin (EC 3.4.21 |
| 18 | 128.5 | 9.2 | 434 | 1 A35005 | u-plasminogen acti |
| 19 | 128.5 | 9.2 | 434 | 1 A35005 | u-plasminogen acti |
| 20 | 126.2 | 9.1 | 433 | 1 A35029 | hepatocyte growth |
| 21 | 125.2 | 9.1 | 433 | 1 A35029 | hepatocyte growth |
| 22 | 124.5 | 8.9 | 810 | 2 B3098 | plasmin (EC 3.4.21 |
| 23 | 124 | 8.9 | 810 | 2 B3098 | plasmin (EC 3.4.21 |
| 24 | 124 | 8.9 | 728 | 1 TRHU | hepatocyte growth |
| 25 | 124 | 8.9 | 728 | 1 TRHU | hepatocyte growth |
| 26 | 123 | 8.8 | 812 | 1 PRBO | plasmin (EC 3.4.21 |
| 27 | 122.5 | 8.8 | 593 | 2 S45281 | coagulation factor |
| 28 | 122 | 8.8 | 4548 | 1 S00657 | apoptosis factor |
| 29 | 121.5 | 8.7 | 603 | 2 S28941 | coagulation factor |
| | | | 728 | 1 A35644 | hepatocyte growth |

ALIGNMENTS

RESULT 1

UKHUT

t-plasminogen activator (EC 3.4.21.66) precursor [validated] - human

N/Alternate names: t-PA, tissue plasminogen activator

C/Species: Homo sapiens (man)

C/Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text change 08-Dec-2000

C/Accession: A94004; A23529; U70562; A93293; S02125; A91343; A93951; A91322; A54645; I60

R/Ref: Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984

A/Title: The structure of the human tissue-type plasminogen activator gene: correlation

A/Accession: A94004; A94004; M01D:842817; PMID:6089198

A/Molecule type: DNA

A/Residues: 1-562 <NT>

A/Cross-references: GB:U00141

A/Note: The codon given for residue 93 (ACC) is inconsistent with the authors' translation

R/Pfizer Degen, S.J.; Rajput, B.; Reich, E.

J. Biol. Chem. 261, 6972-6985, 1986

A/Title: The human tissue plasminogen activator gene.

A/Reference number: A23529; M01D:8616143; PMID:3009482

A/Accession: A23529

A/Molecule type: DNA

A/Residues: 1-562 <DNA>

A/Cross-references: GB:R03021; NID:G339817; P1DN:AAA98809.1; P1D:G339818

R/Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuoka, S.; Higashio, K.

Aggic. Biol. Chem. 55, 1245-1252, 1991

A/Title: Purification and characterization of tissue plasminogen activator secreted by

A/Reference number: U70562; M01D:91291340; PMID:1368681

A/Molecule type: mRNA

A/Residues: 31-562 <TTA>

A/Cross-references: DDBJ:D01096; NID:G220128; P1DN:BA00881.1; P1D:G441174

A/Experimental source: embryonic lung fibroblast IMR-90 cells

A/Note: Part of this sequence, including the amino end of the mature protein, was conf.

R/Pennick, D.; Holmes, M.E.; Kohr, W.J.; Hartings, R.N.; Vohar, G.A.; Ward, C.A.; Bennett

Nature 301, 214-221, 1993

A/Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Escl

A/Reference number: A93293; M01D:8315262; PMID:6337343

A/Accession: A93293

A/Molecule type: mRNA

A/Residues: 1-562 <PEN>

A/Cross-references: GB:U00141

A/Experimental source: melanoma cells

R/Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.

Nucleic Acids Res. 16, 5655, 1988

A/Reference number: 502125; M01D:8626579; PMID:3133640

A/Accession: S02125

A/Status: translation not shown

A/Molecule type: mRNA

A/Residues: 1-562 <SAS>

A/Cross-references: EMBL:X07933; NID:G37243; P1DN:CAA30302.1; P1D:G37244

A:Experimental source: fetal lung cells
R:Kiglatant, H.; Tegawa, H.; Matsunaga, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Matsuda, H.; Teraoka, H.
PMBS Lett. 189, 145-149, 1985
A:Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen activator gene
A:Reference number: A91343; PMID:85285620; PMID:13969853
A:Molecule type: mRNA
A:Accession: A91343
A:Residues: 159, 'G', 86-413, 'E', 435-562 <RXG>
A:Experimental source: Detroit 562 cells; ATCC 138
R:Edlund, T.; Ny, T.; Rånby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
A:Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activator
A:Reference number: A93951; PMID:8316956; PMID:6572897
A:Accession: A93951
A:Molecule type: mRNA
A:Residues: 251-358 <EDL>
A:Experimental source: melanoma cells
R:Pol, G.; Kallstrom, M.; Bergsdott, N.; Wallen, P.; Jorntvall, H.
J. Biol. Chem. 263, 3701-3707, 1984
A:Title: Isolation of human plasminogen activator: peptide analyses confirm an indirectly derived amino acid sequence
A:Reference number: A90489; PMID:85000468; PMID:6433976
A:Contents: annotation; melanoma cells, partial sequence of residues 36-562, active and inactive
R:Pol, G.; Kaplan, L.; Elmarsson, M.; Wallen, P.; Jorntvall, H.
PMBS Lett. 168, 29-32, 1984
A:Title: Differences between urokinase and melanoma forms of tissue plasminogen activator
A:Reference number: A91322; PMID:8418956; PMID:6538514
A:Accession: A91322
A:Molecule type: protein
A:Residues: 33-45;311-320 <PQH>
A:Experimental source: uterus
A:Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln
R:van Zonneveld, A.J.; Verma, H.; Pamekoek, H.
J. Biol. Chem. 263, 4368-4371, 1988
A:Reference number: A37567; PMID:8703961; PMID:3021732
A:Contents: annotation; fibrin binding site
R:van Zonneveld, A.J.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Bouwels, P.H.; Engelen, J.H.; Verma, H.
PMBS Lett. 189, 145-149, 1985
A:Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen activator in fibrinolytic activity
A:Reference number: A37568; PMID:87161761; PMID:3030730
A:Contents: annotation; fibrin binding site
R:Thromb. Haemost. 59, 523-528, 1988
A:Title: Isolation, identification and pharmacokinetic properties of human tissue-type plasminogen activator
A:Reference number: A60902; PMID:8904461; PMID:3142086
A:Contents: annotation; novel form of expressed recombinant t-PA
R:Wille, J.; Loefer, R.; Pahl, T.; Barabon, F.A.O.; Laitle, S.; Smirge, J.S.; Opdenacker, G.; Mollnes, J.E.
PMBS Lett. 189, 145-149, 1985
A:Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its expression
A:Reference number: A54645; PMID:86284200; PMID:3050401
A:Accession: A54645
A:Molecule type: mRNA
A:Residues: 1-562 <HAR>
A:Cross-references: CB:M1518; NID:g190031; PIND:AAA60111.1; PID:g190032
A:Note: parts of this sequence were confirmed by peptide sequencing
R:Reddy, V.B.; Garzamore, A.J.; Sasek, H.; Wei, C.
DNA 6, 461-472, 1987
A:Title: Expression of human uterine tissue-type plasminogen activator in mouse cells
A:Reference number: I60110; PMID:88054470; PMID:2824147
A:Accession: I60110
A:Contents: translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 1-562 <HAR>
A:Cross-references: CB:M18182; NID:g340176; PIND:AAA36000.1; PID:g340177
R:Pislar, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.
J. Biol. Chem. 260, 11223-11230, 1985
A:Title: Isolation and characterization of the human tissue-type plasminogen activator gene
A:Reference number: I55232; PMID:85289338; PMID:3161893
A:Accession: I55232
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-36 <RB2>
A:Cross-references: CB:M1890; NID:g339837; PIND:AAA61213.1; PID:g3398393

C:Comment: Cleavage by plasmin or trypsin produces two chains held together by a salt
 C:Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond
 C:Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat
 C:Genetics:
 A:Gene: GDB:PLAT
 A:Cross-references: GDB:119496; OMIM:173370
 A:Map position: 6p22-6p23
 A:Map coordinates: 124,931,350/1,850,000; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3;
 A:Structure: 124/3; 350/1; 850/0
 A:Keywords: fibrinolytic; glycoprotein; hydrolase; kringle; plasmin; serine proteinase
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-23/Domain: propeptide #status predicted <PRO>
 F:33-562/Product: t-plasminogen activator #status experimental <MAT>
 F:33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
 F:41-78/Domain: fibronectin type I repeat homology <1F>
 F:86-119/Domain: EGF homology <EGF>
 F:127-208/Domain: kringle homology <KR2>
 F:215-96/Domain: kringle homology <KR2>
 F:311-562/Product: t-plasminogen activator chain B #status experimental <BCH>
 F:311-355/Domain: trypsin homology <TRY>
 F:312-145/Domain: 86-97; 91-108; 110-115; 127-208; 148-190; 179-203; 215-286; 236-278; 267-291; 299-
 F:312/145/Domain: catalytic site (Asn) (Gln) (Gln) #status experimental
 F:312/311/Domain: catalytic site (Arg-71) (plasmin, trypsin) #status experimental
 F:310-311/Cleavage site: Arg-71 (plasmin, trypsin) #status experimental
 F:357-406/Active site: His_Asp #status predicted
 F:513/Active site: Ser #status experimental
 Query Match 11.1%; Score 154.5; DB 1; Length 562;
 Best Local Similarity 39.6%; Pred. No. 3,4e-05;
 Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
 Oy 25. CFMDNGHLYREDQTSFAPGLRCLNLDADQSLAAPS-----GAGNSYCRNPDE 75
 Db 127 CLEQGISRGKSTINASGASCTNN--NSSLAQKPSGSRPDARIGLGNHNYCRNPDR 184
 Oy 76 DPDPQGVYSGAGVPEKRCEDLCQSFETS 106
 Db 185 DSK-PMCYVF-KAGKYSERFCPSAGEENS 213
 RESULT 2
 A:35029
 t-plasminogen activator (EC 3.4.21.68) precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A35029; A31597
 R:Feng, P.; Olson, M.; NY, T.
 J:J Biol. Chem. 265, 2025-2027, 1990
 A:1.10.3. Other accession numbers: the 170 base rat tissue-type plasminogen activator gene. Speed
 A:Reference: 145029; A35029; MUID:50110446; PMID:2105315
 A:Accession: A35029
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-559 <FEN>
 A:Cross-references: GB:J31197; NID:9207429; PIND:AAA42261.1; PID:9207431; GB:J05226
 R:NY, T.; Leonardsson, G.; Heuvel, A.J.W.
 DNA 7, 671-677, 1988
 A:Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activator
 A:Reference number: A31597; MUID:89170114; PMID:3148445
 A:Accession: A31597
 A:Molecule type: mRNA
 A:Residues: 1-379; K, 381-559 <NT>
 A:1.10.3. Other accession numbers: the 9530159. PIND:AAA1812.1; PID:9530160
 C:Superfamily: tissue plasminogen activator. EGF homology, fibronectin type I repeat homology
 C:Keywords: fibrinolytic; glycoprotein; hydrolase; kringle; serine proteinase
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-29/Domain: propeptide #status predicted <PRO>
 F:30-559/Product: t-plasminogen activator #status predicted <MAT>
 F:30-508/Product: t-plasminogen activator chain A #status predicted <ACH>
 F:38-75/Domain: fibronectin type I repeat homology <1F>
 F:83-116/Domain: EGF homology <EGF>
 F:124-205/Domain: kringle homology <KR2>
 F:213-294/Domain: kringle homology <KR2>

F.309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F.309-553/Domain: trypsin homology <TRY>
F.38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-224,234-276,265-289,297-4
F.149,481/Binding site: carbohydrate (asn) (covalent) #status predicted
F.308-309/Cleavage site: Arg-116 (plasmin, trypsin) #status predicted
F.355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 11.1%, Score 154, DB 1, Length 559;
Best Local Similarity 32.4%, Pred No. 3, 7e-05, Indels 26, Gaps 6;
Matches 45, Conservative 14, Mismatches 34, Indels 26, Gaps 6;

QY 25 CPDNGHLYREDQSPAPGLRCNWDASGLASAPVS-----GAGNSYCRNPD 75
DB 124 CFEQGITVYGTWSTENGACINW--NSALSGKPYSAARRNMAIKGLGNHYCRNDR 181
QY 76 DRPGMVCVSGEAGVEKRPEDRCPE-----TTSQALPAFTTEIOASGPG 124
DB 182 DYK-FMCCVF-KAGKYTEPCSTPACPKGPTEDCVGKGVTRGTHSPFT--SKASCLPW 237
QY 125 ADEVQVAPANALPARSEA 143
DB 238 NSMILIGKYTAMRANSQA 256

RESULT 3
A.39941
t-plasminogen activator (EC 3.4.21.68) precursor - mouse

C.Species: Mus musculus (house mouse)
C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C.Species: Mus musculus (house mouse)
C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
R.Kitagawa, N. J. Biol. Chem. 263, 1563-1569, 1988
J. Biol. Chem. 263, 1563-1569, 1988
A.Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA

A.Reference number: A23941; MUID:86087303; PMID:2824848
A.Molecule type: mRNA
A.Accession: A23941

A.Residues: 1-559 <RIC>
A.Cross-references: GB:J03520; NID:9202109; PIDD:AAA40470.1; PID:9202110
R.Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.

Enr. J. Biochem. 224, 863-871, 1994
A.Title: Characterization of the murine plasma fibrinolytic system.

A.Reference number: 848202; MUID:95010076; PMID:7523120
A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>

A.Accession: 848202
A.Molecule type: protein
A.Residues: 309-316 <LIU>

A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>

A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>
C.Species: Homo sapiens (man)
C.Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
R.Kitagawa, N. J. Biol. Chem. 263, 1563-1569, 1988
J. Biol. Chem. 263, 1563-1569, 1988
A.Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA

A.Reference number: 848202; MUID:95010076; PMID:7523120
A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>

A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>
C.Species: Homo sapiens (man)
C.Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
R.Kitagawa, N. J. Biol. Chem. 263, 1563-1569, 1988
J. Biol. Chem. 263, 1563-1569, 1988
A.Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA

A.Reference number: 848202; MUID:95010076; PMID:7523120
A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>

A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>
C.Species: Homo sapiens (man)
C.Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
R.Kitagawa, N. J. Biol. Chem. 263, 1563-1569, 1988
J. Biol. Chem. 263, 1563-1569, 1988
A.Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA

A.Reference number: 848202; MUID:95010076; PMID:7523120
A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>

A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>
C.Species: Homo sapiens (man)
C.Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
R.Kitagawa, N. J. Biol. Chem. 263, 1563-1569, 1988
J. Biol. Chem. 263, 1563-1569, 1988
A.Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA

A.Reference number: 848202; MUID:95010076; PMID:7523120
A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>

A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>
C.Species: Homo sapiens (man)
C.Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
R.Kitagawa, N. J. Biol. Chem. 263, 1563-1569, 1988
J. Biol. Chem. 263, 1563-1569, 1988
A.Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA

A.Reference number: 848202; MUID:95010076; PMID:7523120
A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>

A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>
C.Species: Homo sapiens (man)
C.Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
R.Kitagawa, N. J. Biol. Chem. 263, 1563-1569, 1988
J. Biol. Chem. 263, 1563-1569, 1988
A.Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA

A.Reference number: 848202; MUID:95010076; PMID:7523120
A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>

A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>
C.Species: Homo sapiens (man)
C.Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
R.Kitagawa, N. J. Biol. Chem. 263, 1563-1569, 1988
J. Biol. Chem. 263, 1563-1569, 1988
A.Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA

DB 124 CFEQGITVYGTWSTENGACINW--NSVLSLKPYSAARRNMAIKGLGNHYCRNDR 181
QY 76 DRPGMVCVSGEAGVEKRPEDRCPE-----TTSQALPAFTTEIOASGPG 124
DB 182 DYK-FMCCVF-KAGKYTEPCSTPACPKGPTEDCVGKGVTRGTHSPFT--SKASCLPW 237

RESULT 4
A.46688
heparocyte growth factor activator (EC 3.4.21.-) precursor [validated] - human

C.Species: Homo sapiens (man)
C.Date: 21-Sep-1993 #sequence_revision 25-Aug-1995 #text_change 08-Dec-2000
C.Accession: A46688
R.Miyazawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.
J. Biol. Chem. 268, 10024-10028, 1993
A.Title: Molecular cloning and sequence analysis of the cDNA for a human serine protease
d coagulation factor XII.
A.Reference number: A46688; MUID:93252878; PMID:7683665
A.Molecule type: mRNA
A.Residues: 1-655 <MTY>
A.Cross-references: DDBJ:D14012; NID:9219680; PIDD:BA03113.1; PID:9219681
A.Experimental source: liver (mRNA); serum (protein)
A.Note: sequence extracted from NCBI backbone (NCBI:131227, NCBI:131228)
A.Note: parts of the sequence, including the amino ends of the heavy and light chains, C

Genetics: GDB:HGRAC, HGRB, HGRP
A.Cross-references: GDB:1954514
C.Function: 4p16-4p16

A.Description: activates hepatocyte growth factor by specific proteolytic cleavage
A.Pathway: tissue repair and regeneration
C.Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology
C.Keywords: glycoprotein; hydrolase; kringlike; liver; plasma; serine proteinase
F.3-34/Domain: signal sequence #status predicted <SIG>
F.108-148/Domain: fibronectin type II repeat homology <122>
F.164-197/Domain: EGF homology <EG1>
F.202-237/Domain: fibronectin type I repeat homology <121>
F.245-278/Domain: EGF homology <EG2>

F.286-367/Domain: kringlike homology <KRQ>
F.373-407/Domain: hepatocyte growth factor activator light chain #status experimental
F.408-655/Domain: hepatocyte growth factor activator heavy chain #status experimental
F.408-655/Domain: trypsin homology <TRY>
F.40, 48, 290, 469, 492, 546/Binding site: carbohydrate (asn) (covalent) #status predicted
F.164-175, 169, 186, 197, 202-237, 245-256, 250-267, 269-278, 286-367, 307-345, 358-37
F.447, 497, 598/Active site: His, Asp, Ser #status predicted

Query Match 10.5%, Score 146.5, DB 1, Length 655;
Best Local Similarity 36.3%, Pred No. 0, 0.00019, Indels 19, Gaps 6;
Matches 41, Conservative 10, Mismatches 41, Indels 19, Gaps 6;

QY 25 CPDNGHLYREDQSPAPGLRCNWDASGLASAPVS-----GAGNSYCRNPD 76
DB 286 CFEQGITVYGTWSTENGACINW--NSVLSLKPYSAARRNMAIKGLGNHYCRNDR 344

QY 77 DRPGMVCVSGEAGV---EKRPEDRCPEFTTSQALPAFTTEIOASGPG 122
DB 345 ER-FMCCVF-KAGKYTEPCSTPACPKGPTEDCVGKGVTRGTHSPFT--SKASCLPW 237

RESULT 5
A.138098
t-plasminogen activator precursor, inactive endothelial splice form - human

N.Placental names: tissue plasminogen activator
C.Species: Homo sapiens (man)
C.Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
C.Accession: I38098; S01678
R.Kitagawa, N. J. Biol. Chem. 263, 1563-1569, 1988
J. Biol. Chem. 263, 1563-1569, 1988
A.Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA

A.Reference number: I38098; MUID:90192128; PMID:1969145
A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>

A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>
C.Species: Homo sapiens (man)
C.Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
R.Kitagawa, N. J. Biol. Chem. 263, 1563-1569, 1988
J. Biol. Chem. 263, 1563-1569, 1988
A.Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA

A.Reference number: I38098; MUID:90192128; PMID:1969145
A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>

A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>
C.Species: Homo sapiens (man)
C.Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
R.Kitagawa, N. J. Biol. Chem. 263, 1563-1569, 1988
J. Biol. Chem. 263, 1563-1569, 1988
A.Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA

A.Reference number: I38098; MUID:90192128; PMID:1969145
A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>

A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>
C.Species: Homo sapiens (man)
C.Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
R.Kitagawa, N. J. Biol. Chem. 263, 1563-1569, 1988
J. Biol. Chem. 263, 1563-1569, 1988
A.Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA

A.Reference number: I38098; MUID:90192128; PMID:1969145
A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>

A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>
C.Species: Homo sapiens (man)
C.Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
R.Kitagawa, N. J. Biol. Chem. 263, 1563-1569, 1988
J. Biol. Chem. 263, 1563-1569, 1988
A.Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA

A.Reference number: I38098; MUID:90192128; PMID:1969145
A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>

A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>
C.Species: Homo sapiens (man)
C.Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
R.Kitagawa, N. J. Biol. Chem. 263, 1563-1569, 1988
J. Biol. Chem. 263, 1563-1569, 1988
A.Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA

A.Reference number: I38098; MUID:90192128; PMID:1969145
A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>

A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>
C.Species: Homo sapiens (man)
C.Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
R.Kitagawa, N. J. Biol. Chem. 263, 1563-1569, 1988
J. Biol. Chem. 263, 1563-1569, 1988
A.Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA

A.Reference number: I38098; MUID:90192128; PMID:1969145
A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>

A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>
C.Species: Homo sapiens (man)
C.Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
R.Kitagawa, N. J. Biol. Chem. 263, 1563-1569, 1988
J. Biol. Chem. 263, 1563-1569, 1988
A.Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA

A.Reference number: I38098; MUID:90192128; PMID:1969145
A.Molecule type: protein
A.Residues: 33-37, 'X', '39-40 <LIU>

A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-291 <SIE>
A:Cross-references: EMBL:X13097; NID:G35282; PIDD:CAA31489.1; PID:G35283
C:Comment: For the main splice form, see FIR:UKHUT. This form probably does not have pro
C:Genetics:
A:Gene: GDB:PLAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-8p12
A:Insertions: 24/3, 39/1, 85/1, 122/1, 180/2, 211/1, 268/2
C:Superfamily: class plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: alternative splicing; fibrinolysis; glycoprotein; kringle
F:1-23/Domain: signal sequence #status predicted <PRO>
F:23-26/Domain: propeptide #status predicted <PRO>
F:26-32/Domain: plasminogen activator; EGF homology
F:31-32/Domain: fibronectin type III repeat
F:33-36/Domain: EGF homology
F:36-119/Domain: EGF homology
F:127-208/Domain: kringle homology
F:215-291/Domain: kringle homology #status atypical <KR1>
F:41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203/Dissulfide bonds: #status pred
Query Match 10.4%; Score 145.5; DB 2; Length 291;
Best Local Similarity 38.5%; Pred. No. 9.1e-05;
Matches 35; Conservative 6; Mismatches 37; Indels 13; Gaps 4;
QY 25 CFWDNQHLYREDQTSPPAQGLQLMLDAGSLASAPV-----GAGNHSYGRNDE 75
DB 127 CYDDGSIYRGTSWSTAESGACRTNW--NSSALMAQNAVSGRRDARIGLGNHNYCRNPD 184
QY 76 DPRGPMCVYSGEAGVEKRPEDLRCPETTS 106
DB 185 DSR-FWCTVP-KAKGTSSEFCSTPACSEGNS 213
RESULT 6
PMTU
plasmin (EC 3.4.21.7) precursor [validated] - human
N:Alternate names: plasminogen precursor [unreviewed]
N:Contains: angiotensin; microplasma; plasminogen
C:Species: Homo sapiens (man)
C:Date: 24-Apr-1994 #sequence revision 02-Dec-1994 #text change 15-Sep-2000
A:Accession: A35229; 152242; A26646; 162738; 184609; 803735; A00929; A04627; A04625; A04
J:Peteren, T.E.; Martzen, M.R.; Ichinose, A.; Davie, E.W.
J:Biochem. Chem. 265, 6104-6111, 1990
A:Title: Characterization of the gene for human plasminogen, a key proenzyme in the fibr
A:Reference number: A35229; MUID:90202879; PMID:2318848
A:Accession: A35229
A:Molecule type: DNA
A:Residues: 1-810 <PRT>
A:Cross-references: GB:05286; GB:M34276; NID:G19064; PIDD:AAA60113.1; PID:G387026
A:Experimental source: Leukoocyte; Lung Fibroblast
A:Biological source: Human
A:Biological source: Human
A:Biological source: Human
A:Title: Definition of the plasminogen gene in the
A:Reference number: 152242; MUID:91097523; PMID:2266308
A:Accession: 152242
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-16 <MAL1>
A:Cross-references: GB:M6289; NID:G19092; PIDD:AAA36454.1; PID:G35363
R:Forstgen, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.
FBS Lett. 213, 254-260, 1987
A:Title: Molecular cloning and characterization of a full-length cDNA clone for human pl
A:Reference number: A26646; MUID:87162490; PMID:3030813
A:Accession: A26646
A:Molecule type: mRNA
A:Residues: 1-471, 'D', 473-810 <FOR>
A:Cross-references: GB:X05199; NID:G35530; PIDD:CAA28831.1; PID:G35531
A:Experimental source: Liver
A:Biological source: Human
A:Biological source: Human
A:Biological source: Human
A:Title: Cloning of a complementary deoxyribonucleic acid coding for human and
A:Reference number: 145961; MUID:0502311; PMID:6149661
A:Accession: 162738
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 292-471, 'D', 473-810 <MAL2>
A:Cross-references: GB:X02922; NID:G190112; PIDD:AAA60124.1; PID:G387031
A:Accession: 184609
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 367-419 <MAL3>
A:Cross-references: GB:X02921; NID:G190110; PIDD:AAA60123.1; PID:G190111
R:Brinholts, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Leijer, W.; Manneberg, M
Eur. J. Biochem. 114, 465-470, 1981
A:Title: Comparison of the primary structure of the N-terminal CNBR fragments of human,
A:Reference number: 503735; MUID:81212097; PMID:7238497
A:Accession: 503735
A:Molecule type: protein
A:Residues: 20-71, 'E', 73-76 <BRU>
A:Cross-references: GB:X02921; NID:G190110; PIDD:AAA60123.1; PID:G190111
submitted to the Atlas of Protein Sequences and Properties
A:Reference number: A00929
A:Accession: A00929
A:Molecule type: protein
A:Residues: 20-71, 'E', 73-85, 87-106, 'D', 108-360, 'E', 362-810 <SOT>
R:Wiman, B.
Eur. J. Biochem. 76, 129-137, 1977
A:Title: Primary structure of the B-chain of human plasmin.
A:Reference number: A04627; MUID:77225245; PMID:142009
A:Accession: A04627
A:Molecule type: protein
A:Residues: 581-810 <WII>
R:Wiman, B.; Wallen, P.
Eur. J. Biochem. 50, 489-494, 1975
A:Title: Structural relationship between "glutamic acid" and "lysine" forms of human p
A:Reference number: A04625; MUID:75093329; PMID:122932
A:Accession: A04625
A:Molecule type: protein
A:Residues: 20-50, 'Q', 51-71, 'E', 73-85, 87-100 <W12>
R:Wiman, B.; Wallen, P.
Eur. J. Biochem. 58, 539-547, 1975
A:Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen ch
A:Reference number: A04626; MUID:76043692; PMID:126863
A:Accession: A04626
A:Molecule type: protein
A:Residues: 483-507, 'E', 509-604 <W13>
R:Robbins, K.C.; Bernabe, P.; Arzdon, L.; Summarta, L.
J. Biol. Chem. 248, 1631-1633, 1973
A:Title: The primary structure of human plasminogen. II. The histidine loop of human p
A:Reference number: A92125; MUID:73149248; PMID:4694729
A:Contents: annotation; active site
A:Contents: annotation; active site
J:Biochem. Chem. 244, 3550-3597, 1969
A:Title: Studies on the active center of human plasmin. Partial amino acid sequence of
A:Reference number: A92046; MUID:69234739; PMID:4240117
A:Contents: annotation; active site
J:Biochem. Chem. 257, 7401-7406, 1982
A:Title: Structure of the omega-amino-carboxylic acid binding sites of human plasminogen
A:Reference number: A92182; MUID:82213905; PMID:6919539
A:Contents: annotation; omega-amino-carboxylic acid binding sites
R:Wall, Z.; Patches, L.
J. Biol. Chem. 259, 13690-13694, 1984
A:Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential
A:Contents: annotation; fibrin binding site; omega-amino-carboxylic acid binding site
R:Caio, Y.; Yi, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehndel, S.; McCance, S.O.
J. Biol. Chem. 271, 29461-29467, 1996
A:Title: Kringle domains of human angiotensin. Characterization of the anti-proliferative
A:Reference number: A58811; MUID:97067211; PMID:8910613
A:Contents: annotation
R:Jinjen, H.R.; Ugwu, F.; Banl, A.; Collen, D.
Biochemistry 37, 4699-4702, 1998
A:Title: Generation of an angiotensin-like fragment from plasminogen by stromelysin-1 (

A:Reference number: A58812; PMID:9548733; PMID:9548733
A:Contents: annotation
R:Mulichak, A.M.
Submitted to the Brookhaven Protein Data Bank, July 1991
A:Reference number: A51341; PDB:1PK4
A:Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454
R:Tullinsky, A.; Wu, T.P.
Submitted to the Brookhaven Protein Data Bank, July 1991
A:Reference number: A51498; PDB:2PK4
A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 375-454
R:Wu, T.P.; Tullinsky, A.
Submitted to the Brookhaven Protein Data Bank, August 1993
A:Reference number: A51911; PDB:1PKR
A:Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-181
R:Padmanabhan, K.; Tullinsky, A.
Submitted to the Brookhaven Protein Data Bank, April 1994
A:Reference number: A52408; PDB:1PMK
A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-454
R:Tullinsky, A.; Mathews, I.I.
Submitted to the Brookhaven Protein Data Bank, December 1995
A:Reference number: A65324; PDB:1CEA
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R:Tullinsky, A.; Mathews, I.I.
Submitted to the Brookhaven Protein Data Bank, December 1995
A:Reference number: A65325; PDB:1CEB
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R:Mulichak, A.M.; Tullinsky, A.; Ravichandran, K.G.
Biochemistry 30, 10576-10588, 1991
A:Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.9 Å
A:Reference number: A58819; PMID:92031502; PMID:1657148
A:Contents: annotation
R:Wu, T.P.; Padmanabhan, K.; Tullinsky, A.; Mulichak, A.M.
Biochemistry 30, 10589-10594, 1991
A:Title: The refined structure of the epsilon-aminocaproic acid complex of human plasminogen kringle 4 refined at 1.9 Å
A:Reference number: A58818; PMID:92031503; PMID:1657149
A:Contents: annotation
R:De Vos, A.M.; Ulrich, M.H.; Kelley, R.F.; Padmanabhan, K.; Tullinsky, A.; Westbrook, M.
Biochemistry 31, 270-279, 1992
A:Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator at 2.2 Å
A:Reference number: A53483; PMID:9211803; PMID:1120013
A:Contents: annotation; X-ray crystallography, 2.4 angstroms
R:Seec, A.; Tellez, M.; Winkler, M.; Yamano, A.
Submitted to Brookhaven Protein Data Bank, June 1995
A:Reference number: A65980; PDB:1KRN
A:Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-454
R:Rejzender, M.; Llinas, M.
Submitted to the Brookhaven Protein Data Bank, August 1996
A:Reference number: A65803; PDB:1HPJ
A:Contents: annotation
R:Rejzender, M.; Llinas, M.
Submitted to the Brookhaven Protein Data Bank, August 1996
A:Reference number: A65804; PDB:1HPK
A:Contents: annotation; conformation by (1)H-NMR, residues 103-181
R:Rejzender, M.R.; Llinas, M.
Submitted to the Brookhaven Protein Data Bank, August 1996
A:Reference number: A54645; PMID:94237157; PMID:8181475
A:Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 1.
A:Contents: annotation; conformation by (1)H-NMR, residues 96-184
R:Rejzender, M.R.; Llinas, M.
Submitted to the Brookhaven Protein Data Bank, August 1996
A:Title: Solution structure of the epsilon-aminocaproic acid complex of human plasminogen kringle 1.
A:Reference number: A58817; PMID:94237158; PMID:8181476
A:Contents: annotation; conformation by (1)H-NMR
R:Comment: plasminogen is synthesized by the kidney and is present in plasma and many other tissues
C:Comment: plasminogen is converted to plasmin by plasminogen activators (see PIR:UKTU & PIR:GHNGB)
C:Comment: Plasmin is inactivated by alpha-2-antiplasmin (see PIR:ITVUA2) immediately after its release from the fibrinogen-actin complex
C:Comment: Microplasmin is formed by autocatalytic cleavage of plasmin under artificial conditions
C:Comment: Streptelysin I (see PIR:KCHUS1) acts on plasminogen to produce angiotensin. This is a major component of the renin-angiotensin system
A:Gene: GDB:PLG
A:Genetics:

A:Cross-references: GDB:119498; OMIM:113350
A:Map position: 6q26-qc27
A:Intron: 17/1; 62/2; 98/1; 136/2; 183/1; 223/2; 263/1; 317/2; 366/1; 419/2; 480/1; 523/2
C:Function:
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a variety of tissues; cleaves the walls of the grana of chloroplasts; also activates the urokinase-type plasminogen activator
A:Pathway: fibrinolysis
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: angiotensin inhibitor; blood duplication; fibrinolysis; glycoprotein; hydrolase
F:1-96/Domain: plasminogen-related protein precursor homology <PLH>
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-910/Product: plasminogen #status experimental <PRO>
F:20-96/Domain: activation peptide #status experimental <APT>
F:79-466/Product: angiotensin #status experimental <AST>
F:97-580/Domain: plasmin chain A #status experimental <CHA>
F:103-181/Domain: kringle homology <KR1>
F:185-262/Domain: kringle homology <KR2>
F:215-352/Domain: kringle homology <KR3>
F:317-454/Domain: kringle homology <KR4>
F:481-560/Domain: kringle homology <KR5>
F:550-580,581-810/Product: microplasmin #status experimental <MMT>

Query Match 10.2%; Score 142; DB 1; Length 810;
Best Local Similarity 37.2%; Pred. No. 0.00058; Indels 12; Gaps 4;
Matches 32; Conservative 8; Mismatches 34;

Ox 25 CFWNDGHLTYREDOTSPADGLRCLNLDLDA-----QSGLASAPVSGAGNHSYCRNDDPDPG 79
Db 103 CKTGGKNGKNGYTKMSTKTKNGITCKQKMSSTSPHPRSPRSPHNSGL-ENNYCRNDDNDPDS 161
Ox 80 PWCYVSGEAGVPEKR--PCEDIRCPCE 103
Db 162 PWCITTD----PKRKYDCILCECE 183

RESULT 7
A:Accession: A40522
A:Plasma (EC 3.4.21.7) precursor - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Feb-1992 #sequence_revision 17-Apr-1993 #ext_change 16-Jul-1999
C:Accession: A40522
R:Kanaias, J.J.; Makker, S.P.
J: 8101; Chem. 266, 10825-10829, 1991
A:Title: Identification of the rat Heymann nephritis autoantigen (GP330) as a receptor
A:Reference number: A40522; MUID:91250378; PMID:1645711
A:Accession: A40522
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-169 <RXN>
A:Cross-references: GB:M62823; NID:926215; PIDN:AAA1884.1; PID:G554488
A:Note: the authors translated the codon TCT for residue 76 as Ala
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:34-112/Domain: kringle homology <KR>
F:34-112,55-95,83-107/Disulfide bonds: #status predicted

Query Match 9.9%; Score 138; DB 2; Length 169;
Best Local Similarity 31.8%; Pred. No. 0.00021; Indels 14; Gaps 5;
Matches 35; Conservative 13; Mismatches 48;

Ox 25 CFWNDGHLTYREDOTSPADGLRCLNLDLDA-----QSGLASAPVSGAGNHSYCRNDDPDPG 79
Db 34 CYQNGKNGKYSTSTNTTKCKQSGVSMTPHSHSKTPANFPDGL-ENNYCRNDDNDPDS 92
Ox 80 PWCYVSGEAGVPEKR--PCEDIRCPCEITTSQALPAFTTEICPABSGAGDE 127
Db 93 PWCFTTD---PSVFWCYCNLRKSCETGGGV--ASALIVPQVPAEPGIS 136

RESULT 8
A:Cross-references: GDB:119498; OMIM:113350
A:Map position: 6q26-qc27
A:Intron: 17/1; 62/2; 98/1; 136/2; 183/1; 223/2; 263/1; 317/2; 366/1; 419/2; 480/1; 523/2
C:Function:
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a variety of tissues; cleaves the walls of the grana of chloroplasts; also activates the urokinase-type plasminogen activator
A:Pathway: fibrinolysis
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: angiotensin inhibitor; blood duplication; fibrinolysis; glycoprotein; hydrolase
F:1-96/Domain: plasminogen-related protein precursor homology <PLH>
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-910/Product: plasminogen #status experimental <PRO>
F:20-96/Domain: activation peptide #status experimental <APT>
F:79-466/Product: angiotensin #status experimental <AST>
F:97-580/Domain: plasmin chain A #status experimental <CHA>
F:103-181/Domain: kringle homology <KR1>
F:185-262/Domain: kringle homology <KR2>
F:215-352/Domain: kringle homology <KR3>
F:317-454/Domain: kringle homology <KR4>
F:481-560/Domain: kringle homology <KR5>
F:550-580,581-810/Product: microplasmin #status experimental <MMT>

C.Species: Papio cynocephalus, Papio hamadryas cynocephalus (Yellow baboon)
C.Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999
C.Accession: S14687, S08651
R.Au: Y.P.T.; Wang, T.W.; Clower, A.W.
Nucleic Acids Res 18, 3411, 1990
A.Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen
A.Reference number: S14687, MUID:90287734; PMID:2113276
A.Accession: S14687
A.Molecule type: mRNA
A.Residues: 1-433 <AV>
A.Cross-references: EMBL:X51935; NID:938130; PDB:CAA6200.1; PID:938131
C.Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C.Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F.1-20/Domain: signal sequence #status predicted <SIG>
F.21-176/Product: plasminogen activator chain A #status predicted <ACH>
F.30-61/Domain: EGF homology <EGF>
F.69-150/Domain: kringle homology <KRG>
F.178-433/Product: plasminogen activator chain B #status predicted <BCH>
F.178-433/Domain: trypsin homology <TRY>
F.167-298,208-224,216-287,315-384,347-363,374-402/Disulfide bonds: #status predicted
F.223,274,378/Active site: His, Asp, Ser #status predicted
F.324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.9%; Score 138; DB 1; Length 433;
Best local similarity 32.7%; Pred. No. 0.00061;
Matches 32; Conservative 14; Mismatches 36; Indels 16; Gaps 4;

Qy 25 CFMDNGHLYREDQTSPPAGRLCLNWLDA-----OSGLASAPVSGAGNHSYCRNPEDP 77
Db 69 CYGNGHFRGKASTDTMGHSLAVNSATVLTQYHAHSDALDGLGKHNCRNPD-NR 127
Qy 78 RGPWCYVSGEAGVPEK-----RPECEDLRCPETTSQAL 109
Db 128 RRPWCYV--QVGLKRVQECVWNCADKCKSSPPEEL 163

RESULT 9
UCS061
C.Species: Rattus norvegicus (Norway rat)
C.Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 16-Jun-2000
C.Accession: UCS061
R.Ohshima, K.; Iwama, A.; Matsuno, K.; Esaki, T.; Sakamoto, O.; Hamaguchi, I.; Takasu, N
Biochem. Biophys. Res. Commun. 227, 273-280, 1996
A.Title: Molecular cloning of rat urokinase-type plasminogen activating protein and its involvement in
A.Reference number: UCS061; MUID:97011126; PMID:8858136
A.Molecule type: mRNA
A.Residues: 1-716 <CH>
A.Cross-references: EMBL:X5096; NID:91669718; PDB:CAA64473.1; PID:91669719
C.Superfamily: disulfide-bonded heterodimer of chains derived from the same precursor
C.Keywords: duplication; glycoprotein; growth factor; kringle
F.1-31/Domain: signal sequence #status predicted <SIG>
F.32-488,489-716/Product: macrophage-stimulating protein 1 #status predicted <MAT>
F.130-186/Domain: macrophage-stimulating protein 1 alpha chain #status predicted <ACH>
F.191-268/Domain: kringle homology <KR12>
F.292-370/Domain: kringle homology <KR13>
F.379-457/Domain: kringle homology <KR14>
F.489-716/Domain: macrophage-stimulating protein 1 beta chain #status predicted <BCH>
F.489-709/Domain: trypsin homology <TRY>
F.72,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.8%; Score 137; DB 1; Length 716;
Best local similarity 27.3%; Pred. No. 0.0013;
Matches 48; Conservative 12; Mismatches 56; Indels 60; Gaps 8;

Qy 2 LLAVQAFLVSNMLLAELVSG-----GCFMDNGHLYREDQTSPPAGRLCLNWL 49
Db 80 LLPWTO-----HSLAQLHSHSLCDLFOKQDYATCTMGNCASRGVARTADGLPCQAW 134
Qy 50 ---LDAOSGLASAPVSGAGNHSYCRNPEDPGRWCYV----- 85

Db 135 SRAFPNDKHYPTKNGLT-ENFCRNPDPGRPWCYVTSNVSVPFSGCISKEAVCVW 193
Qy 86 -----GGAIVPEK-RPCE--DLRCPET-----TSQALPAFTTEIOENSGP 123
Db 194 CNGEDYRGSVDTESGREGCQRMWDLQHPHSHRPFPEKPDKALKNYCRNPASERP 249

RESULT 10
UKRG
C.Species: Sus scrofa domestica (domestic pig)
C.Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
C.Accession: A00932
R.Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
Nucleic Acids Res. 12, 9525-9541, 1984
A.Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
A.Reference number: A00932; MUID:85087954; PMID:6096832
A.Accession: A00932
A.Molecule type: DNA
A.Residues: 1-240, 'H', 242-442 <NAG1>
A.Experimental source: kidney cell line LLC-PK1
R.Nagamine, Y.
Submitted to the Protein Sequence Database, December 1986
A.Reference number: A37566
A.Contents: annotation; correction to residue 241
C.Intentions: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 268/1; 335/1; 384/3
C.Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology.
C.Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F.1-20/Domain: signal sequence #status predicted <SIG>
F.21-186/Product: urokinase-type plasminogen activator chain A #status predicted <AC
F.23-64/Domain: EGF homology <EGF>
F.72-153/Domain: kringle homology <KRG>
F.180-442/Product: urokinase-type plasminogen activator chain B #status predicted <B
F.180-442/Domain: trypsin homology <TRY>
F.152,150,429,236,228-229,323-353,356-372,383-411/Disulfide bonds: #status predicted
F.433,286,387/Active site: His, Asp, Ser #status predicted

Query Match 9.7%; Score 135.5; DB 1; Length 442;
Best local similarity 36.9%; Pred. No. 0.001;
Matches 38; Conservative 7; Mismatches 37; Indels 21; Gaps 5;

Qy 25 CFMDNGHLYREDQTSPPAGRLCLNWLDAQSL-----ASAPVS---GAGNHSYCRNPEDP 77
Db 72 CFEGNGHFRGKASTDTMGHSLAVNSATVLTQYHAHSDALDGLGKHNCRNPD-NQ 130
Qy 78 RGPWCYVSGEAGVPEK-----RPECEDLRCPETTSQ 107
Db 131 RRPWCYV--QVGLKRVQECVWNCASGESHRAVDGKNPSTPE 173

RESULT 11
JS0599
C.Species: Desmodus rotundus (common vampire bat)
C.Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C.Accession: JS0599
R.Kratzschmar, J.; Heindler, B.; Langer, G.; Bojdel, W.; Brlingmann, P.; Alagon, A.; Do
Gene 105, 229-237, 1991
A.Title: The plasminogen activator family from the salivary gland of the vampire bat De
A.Reference number: JS0597; MUID:92039036; PMID:1937019
A.Accession: JS0599
A.Molecule type: mRNA
A.Residues: 1-431 <KRA>
A.Cross-references: GB:M63989; NID:9166076; PDB:AAA31594.1; PID:9166077
C.Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat ho
C.Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F.1-21/Domain: signal sequence #status predicted <SIG>
F.22-36/Domain: propeptide #status predicted <PRO>


```

Db      80  LIPNQ-----HSIAHLOVHSLICLIFOKDYRITCIMNGSVSYGVARTAGLPGQAM 134
Qy      50  ---LDAQGLSAVASQAGNHSCYCPNPPDPGMCYVS-----85
Db      135  SRPFPNDHKYTPPKQGL-EENFCNPDPGDPGPMCYTTNNSVAFQSCGKTCREAVCVL 193
Qy      86  -----GEAVPER-KPCE--DLCAPT 104
Db      194  CNGEDYRGEVDVYTSGRGECRDLDHPHS 222

```

RESULT 15

```

N:Plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #ext_change 16-Jul-1999
C:Accession: J06600
R:Kraetzschmar, J.; Hendler, B.; Langer, G.; Beidol, W.; Bringmann, P.; Alagon, A.; Dor
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat
A:Reference number: J059577; PMID:92039036; PMID:11937019
A:Accession: J06060
A:Molecule type: mRNA
A:Residues: 1-394 <KR>
A:Cross-references: GB:M63990; NID:J166078; PDB:AAA1555.1; PID:J166079
A:Note: the authors translated the codon ATC for residue 75 as Thr
C:Keyword: fibinolytic; tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
F:1-21/Domain: signal sequence #status predicted <sig>
F:12-36/Domain: propeptide #status predicted <PRO>
F:14-134/Domain: single homology <XRG>
F:14-136/Domain: cystein homology <TRY>
F:14-126-66-108-97-121-131-262-11-190-189-251-276-351-308-324-341-369/Disulfide bonds:
F:142-143/Cleavage site: His-Ser (plasma) #status predicted
F:189-239-346/Active site: His, Asp, Ser #status predicted
F:135/Binding site: carbohydrate (Am) (covalent) #status predicted

Query Match          9 54;  Score 132;  DB 2;  Length 394;
Best Local Similarity 32.74;  Pred: No. 0.0018;
Matches 33;  Conservative 9;  Mismatches 39;  Indels 20;  Gaps 5;

Oy 17 AENYGS--GGCWNDGHLHYREDQTSFAPGLRLIN-----LDAQSGLASAPVSGAGNH 67
Db 35 SRATGSPATYCKDGGVYTKRWSTSSGAGQCLINMNSMLIRRTYNGRPEAVKGLGNH 94
Oy 68 SYRANDEDPRGQPCVY-----SGAGVPEKRCECDLRC 101
Db 95 NYCRNPDGASK-RPWCYTKARKFTSSCSVP---VCSFATC 131

```

Search completed: April 7, 2003, 07:00:44
Job time : 20.6747 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 06:30:08 ; Search time 9.89785 Seconds

(without alignments)
1102.085 Million cell updates/sec

Title: US-10-057-951-2

Perfect score: 1393

Sequence: 1 MLAWQAFVSNMLAEAY.....PVDPGSGPTLMQAGTPGA 263

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | * Query Match | Length | ID | Description |
|------------|-------|---------------|--------|--------------|--------------------|
| 1 | 160.5 | 11.5 | 566 | 1 TPA_BOVIN | Q28198 bos taurus |
| 2 | 154.5 | 11.1 | 562 | 1 TPA_HUMAN | P00750 homo sapien |
| 3 | 154 | 11.1 | 559 | 1 TPA_RAT | P19637 ratius norv |
| 4 | 150 | 10.8 | 653 | 1 HGF_MOUSE | Q9R098 mus musculu |
| 5 | 147.5 | 10.6 | 559 | 1 TPA_MOUSE | P11214 mus musculu |
| 6 | 146.5 | 10.5 | 655 | 1 HGF_HUMAN | Q04756 homo sapien |
| 7 | 142 | 10.2 | 810 | 1 PLMN_HUMAN | P00747 homo sapien |
| 8 | 138 | 9.9 | 169 | 1 PLMN_RAT | Q01177 ratius norv |
| 9 | 138 | 9.9 | 433 | 1 UROK_PAPCY | P16227 papio cynoc |
| 10 | 135.5 | 9.7 | 442 | 1 UROK_PIG | P04185 sus scrofa |
| 11 | 134 | 9.6 | 431 | 1 URTB_DESRO | P98121 desmodus ro |
| 12 | 134 | 9.6 | 477 | 1 URT1_DESRO | P15638 desmodus ro |
| 13 | 134 | 9.6 | 716 | 1 HGF_MOUSE | P26828 mus musculu |
| 14 | 132 | 9.5 | 394 | 1 UROK_DESRO | P49150 desmodus ro |
| 15 | 132 | 9.5 | 431 | 1 UROK_HUMAN | P00745 homo sapien |
| 16 | 130 | 9.3 | 810 | 1 PLMN_BRIEU | Q29485 etineus e |
| 17 | 128.5 | 9.2 | 434 | 1 UROK_CHICK | P15120 gallus gall |
| 18 | 127 | 9.1 | 728 | 1 UROK_BOVIN | Q05889 bos taurus |
| 19 | 126.5 | 8.9 | 810 | 1 HGF_MOUSE | Q08048 mus musculu |
| 20 | 124.5 | 8.9 | 728 | 1 THRB_HUMAN | P12454 macaca mula |
| 21 | 124 | 8.9 | 622 | 1 HGF_HUMAN | P00734 homo sapien |
| 22 | 124 | 8.9 | 728 | 1 HGF_MOUSE | P14210 homo sapien |
| 23 | 123 | 8.8 | 812 | 1 PLMN_BOVIN | P06868 bos taurus |
| 24 | 123 | 8.8 | 593 | 1 PLMN_BOVIN | P98149 bos taurus |
| 25 | 122.5 | 8.8 | 4548 | 1 APOA_HUMAN | P08519 homo sapien |
| 26 | 122 | 8.8 | 603 | 1 PA12_CAVPO | Q04862 cavia porce |
| 27 | 121.5 | 8.7 | 728 | 1 HGF_RAT | P17945 ratius norv |
| 28 | 120.5 | 8.7 | 333 | 1 PLMN_CANFA | P80009 canis famli |
| 29 | 120 | 8.6 | 433 | 1 UROK_MOUSE | P06867 sus scrofa |
| 30 | 120 | 8.6 | 790 | 1 PLMN_PIG | P06867 sus scrofa |
| 31 | 119.5 | 8.5 | 477 | 1 URT1_DESRO | P98148 homo sapien |
| 32 | 118.5 | 8.5 | 615 | 1 PA12_HUMAN | P29598 ratius norv |
| 33 | 118 | 8.5 | 432 | 1 UROK_RAT | |

| | | | | | |
|----|-------|-----|------|--------------|--------------------|
| 34 | 117 | 8.4 | 625 | 1 THRB_BOVIN | P00735 bos taurus |
| 35 | 116 | 8.3 | 1420 | 1 APOA_MACMU | P14417 macaca mula |
| 36 | 115.5 | 8.3 | 711 | 1 HGF_HUMAN | P26927 homo sapien |
| 37 | 114.5 | 8.2 | 812 | 1 PLMN_MOUSE | P26918 mus musculu |
| 38 | 114 | 8.2 | 473 | 1 KREM_MOUSE | Q29483 mus musculu |
| 39 | 114 | 8.2 | 473 | 1 KREM_RAT | Q22484 ratius norv |
| 40 | 114 | 8.2 | 475 | 1 KREM_HUMAN | Q26484 homo sapien |
| 41 | 114 | 8.2 | 618 | 1 THRB_MOUSE | P19221 mus musculu |
| 42 | 109.5 | 7.9 | 1709 | 1 SN_HUMAN | Q26222 homo sapien |
| 43 | 107 | 7.7 | 617 | 1 THRB_RAT | P18292 ratius norv |
| 44 | 103.5 | 7.4 | 343 | 1 PLMN_SHEEP | P81286 ovis aries |
| 45 | 101.5 | 7.3 | 325 | 1 PLMN_PETMA | P33574 petromyzon |

ALIGNMENTS

```

RESULT 1
ID      TPA_BOVIN          STANDARD;          PRT;          566 AA.
AC      Q28198
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE      (t-PA) (t-plasminogen activator).
GN      Bos taurus (Bovine).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC      Bovidae; Bovinae; Bos.
OX      NCBI_TaxID=9913;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      Tissue-Kidney;
RA      Rayn P. Berglund L., Petersen T.E.;
RT      "Cloning and characterization of the bovine plasminogen activators uPA
RT      and tPA.";
RL      Int. Dairy J. 5:605-617(1995).
CC      -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
CC      TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC      CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC      ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC      MANY OTHER PHYSIOPATHOLOGICAL EVENTS.
CC      -1- CATALYTIC ACTIVITY: Specific cleavage of Arg|-Val bond in
CC      plasminogen to form plasmin.
CC      -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC      BOND.
CC      -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC      -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC      PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC      ARG-314 CATALYZED BY PLASMIN. TISSUE KALLIKREIN OR FACTOR XA.
CC      -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC      CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC      -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC      -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC      -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE 1 DOMAIN.
CC      -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC      -----
CC      This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC      or send an email to license@ebi.ac.uk).
CC      -----
DR      EMBL; X65800; CAAS9795.1; -.
DR      HSSP; P00750; 1RTF.
DR      MEROPS; S01.232; -.
DR      InterPro; IPR001314; Chymotrypsin.
DR      InterPro; IPR000561; EGF-like.
DR      InterPro; IPR000083; Fibrinectn.

```


| Query Match | Best local similarity | Score | Length | DB 1 | DB 2 | DB 3 | DB 4 | DB 5 | DB 6 | DB 7 | DB 8 | DB 9 | DB 10 | DB 11 | DB 12 | DB 13 | DB 14 | DB 15 | DB 16 | DB 17 | DB 18 | DB 19 | DB 20 | DB 21 | DB 22 | DB 23 | DB 24 | DB 25 | DB 26 | DB 27 | DB 28 | DB 29 | DB 30 | DB 31 | DB 32 | DB 33 | DB 34 | DB 35 | DB 36 | DB 37 | DB 38 | DB 39 | DB 40 | DB 41 | DB 42 | DB 43 | DB 44 | DB 45 | DB 46 | DB 47 | DB 48 | DB 49 | DB 50 | DB 51 | DB 52 | DB 53 | DB 54 | DB 55 | DB 56 | DB 57 | DB 58 | DB 59 | DB 60 | DB 61 | DB 62 | DB 63 | DB 64 | DB 65 | DB 66 | DB 67 | DB 68 | DB 69 | DB 70 | DB 71 | DB 72 | DB 73 | DB 74 | DB 75 | DB 76 | DB 77 | DB 78 | DB 79 | DB 80 | DB 81 | DB 82 | DB 83 | DB 84 | DB 85 | DB 86 | DB 87 | DB 88 | DB 89 | DB 90 | DB 91 | DB 92 | DB 93 | DB 94 | DB 95 | DB 96 | DB 97 | DB 98 | DB 99 | DB 100 | DB 101 | DB 102 | DB 103 | DB 104 | DB 105 | DB 106 | DB 107 | DB 108 | DB 109 | DB 110 | DB 111 | DB 112 | DB 113 | DB 114 | DB 115 | DB 116 | DB 117 | DB 118 | DB 119 | DB 120 | DB 121 | DB 122 | DB 123 | DB 124 | DB 125 | DB 126 | DB 127 | DB 128 | DB 129 | DB 130 | DB 131 | DB 132 | DB 133 | DB 134 | DB 135 | DB 136 | DB 137 | DB 138 | DB 139 | DB 140 | DB 141 | DB 142 | DB 143 | DB 144 | DB 145 | DB 146 | DB 147 | DB 148 | DB 149 | DB 150 | DB 151 | DB 152 | DB 153 | DB 154 | DB 155 | DB 156 | DB 157 | DB 158 | DB 159 | DB 160 | DB 161 | DB 162 | DB 163 | DB 164 | DB 165 | DB 166 | DB 167 | DB 168 | DB 169 | DB 170 | DB 171 | DB 172 | DB 173 | DB 174 | DB 175 | DB 176 | DB 177 | DB 178 | DB 179 | DB 180 | DB 181 | DB 182 | DB 183 | DB 184 | DB 185 | DB 186 | DB 187 | DB 188 | DB 189 | DB 190 | DB 191 | DB 192 | DB 193 | DB 194 | DB 195 | DB 196 | DB 197 | DB 198 | DB 199 | DB 200 | DB 201 | DB 202 | DB 203 | DB 204 | DB 205 | DB 206 | DB 207 | DB 208 | DB 209 | DB 210 | DB 211 | DB 212 | DB 213 | DB 214 | DB 215 | DB 216 | DB 217 | DB 218 | DB 219 | DB 220 | DB 221 | DB 222 | DB 223 | DB 224 | DB 225 | DB 226 | DB 227 | DB 228 | DB 229 | DB 230 | DB 231 | DB 232 | DB 233 | DB 234 | DB 235 | DB 236 | DB 237 | DB 238 | DB 239 | DB 240 | DB 241 | DB 242 | DB 243 | DB 244 | DB 245 | DB 246 | DB 247 | DB 248 | DB 249 | DB 250 | DB 251 | DB 252 | DB 253 | DB 254 | DB 255 | DB 256 | DB 257 | DB 258 | DB 259 | DB 260 | DB 261 | DB 262 | DB 263 | DB 264 | DB 265 | DB 266 | DB 267 | DB 268 | DB 269 | DB 270 | DB 271 | DB 272 | DB 273 | DB 274 | DB 275 | DB 276 | DB 277 | DB 278 | DB 279 | DB 280 | DB 281 | DB 282 | DB 283 | DB 284 | DB 285 | DB 286 | DB 287 | DB 288 | DB 289 | DB 290 | DB 291 | DB 292 | DB 293 | DB 294 | DB 295 | DB 296 | DB 297 | DB 298 | DB 299 | DB 300 | DB 301 | DB 302 | DB 303 | DB 304 | DB 305 | DB 306 | DB 307 | DB 308 | DB 309 | DB 310 | DB 311 | DB 312 | DB 313 | DB 314 | DB 315 | DB 316 | DB 317 | DB 318 | DB 319 | DB 320 | DB 321 | DB 322 | DB 323 | DB 324 | DB 325 | DB 326 | DB 327 | DB 328 | DB 329 | DB 330 | DB 331 | DB 332 | DB 333 | DB 334 | DB 335 | DB 336 | DB 337 | DB 338 | DB 339 | DB 340 | DB 341 | DB 342 | DB 343 | DB 344 | DB 345 | DB 346 | DB 347 | DB 348 | DB 349 | DB 350 | DB 351 | DB 352 | DB 353 | DB 354 | DB 355 | DB 356 | DB 357 | DB 358 | DB 359 | DB 360 | DB 361 | DB 362 | DB 363 | DB 364 | DB 365 | DB 366 | DB 367 | DB 368 | DB 369 | DB 370 | DB 371 | DB 372 | DB 373 | DB 374 | DB 375 | DB 376 | DB 377 | DB 378 | DB 379 | DB 380 | DB 381 | DB 382 | DB 383 | DB 384 | DB 385 | DB 386 | DB 387 | DB 388 | DB 389 | DB 390 | DB 391 | DB 392 | DB 393 | DB 394 | DB 395 | DB 396 | DB 397 | DB 398 | DB 399 | DB 400 | DB 401 | DB 402 | DB 403 | DB 404 | DB 405 | DB 406 | DB 407 | DB 408 | DB 409 | DB 410 | DB 411 | DB 412 | DB 413 | DB 414 | DB 415 | DB 416 | DB 4 |
|-------------|-----------------------|-------|--------|------|------|------|------|------|------|------|------|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|------|
|-------------|-----------------------|-------|--------|------|------|------|------|------|------|------|------|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|------|

| ID | TPA | HUMAN | STANDARD; | PRT; | 562 AA. |
|----|--|-----------------------------------|-----------|------|---------|
| AC | P00750; | Q15103; | | | |
| DT | 21-JUL-1986 | (Rel. 01, Created) | | | |
| DT | 21-JUL-1986 | (Rel. 01, Last sequence update) | | | |
| DT | 15-JUN-2002 | (Rel. 41, Last annotation update) | | | |
| DE | tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA) | | | | |
| DE | (t-Pa) (t-plasminogen activator) (Alteplase) (Retelase). | | | | |
| GN | PLAT. | | | | |
| OS | Homo sapiens (human). | | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | | |
| OX | NCBI_TaxId=9606; | | | | |
| PN | [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RP | TISSUE-Melanoma; | | | | |
| RC | MEDLINE=83115252; PubMed=3137343; | | | | |
| EX | Samuelson D., Holmes M.E., Koh M.J., Harting R.N., Vohar G.A., | | | | |
| RA | Ward C.A., Bennett M.F., Yelverton E., Seeburg P.H., Heynaker H.L., | | | | |
| RA | Geeddel D.V., Colten D.; | | | | |
| RT | "Cloning and expression of human tissue-type plasminogen activator | | | | |
| RT | cDNA in E. coli."; | | | | |
| RL | Nature 301:214-221 (1983). | | | | |
| RN | [2] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RP | TISSUE-Fetal lung; | | | | |
| RC | MEDLINE=88262579; PubMed=3133640; | | | | |
| RX | Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.; | | | | |
| RA | "Nucleotide sequence of the tissue-type plasminogen activator cDNA | | | | |
| RT | from human fetal lung cells."; | | | | |
| RL | Nucleic Acids Res. 15:5695-5695 (1988). | | | | |
| RN | [3] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RX | MEDLINE=88054470; PubMed=2824147; | | | | |
| RA | Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Galli J., | | | | |
| RA | Heinig N.; | | | | |
| RT | "Expression of human urokinase tissue-type plasminogen activator in | | | | |
| RT | mouse cells using BPV vectors." | | | | |
| RL | DNA 6:461-472 (1987). | | | | |
| RN | [4] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RX | MEDLINE=86196143; PubMed=3009482; | | | | |
| RA | Friedner Degen S.J., Rajput B., Reich E.; | | | | |
| RA | "The human tissue plasminogen activator gene."; | | | | |
| RL | J Biol. Chem. 261:6572-6585 (1986). | | | | |
| RN | [5] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RX | MEDLINE=84296137; PubMed=6089199; | | | | |
| RA | "My T., Elgh F., Lund B.; | | | | |
| RT | The structure of the human tissue-type plasminogen activator gene: | | | | |
| RT | correlation of intron and exon structures to functional and | | | | |
| RT | structural domains."; | | | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359 (1984). | | | | |
| RN | [6] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RP | MEDLINE=86284200; PubMed=3090401; | | | | |
| RX | Harris T.J., Patel T., Maxson F.A., Little S., Emtage J.S., | | | | |
| RA | Odenacker G., Volckaert G., Rombaux W., Billiau A., Somer P.; | | | | |
| RT | "Cloning of cDNA coding for human tissue-type plasminogen activator | | | | |
| RT | and its expression in Escherichia coli."; | | | | |
| RL | Mol. Biol. Med. 3:279-292 (1986). | | | | |
| RN | [7] | | | | |
| RP | SEQUENCE OF 212-361 FROM N.A. | | | | |
| RX | MEDLINE=83169656; PubMed=6577889; | | | | |
| RA | Edlund T., Ny T., Rasmussen M., Heben L.-O., Palm G., Holmgren E., | | | | |
| RA | Josephson S. | | | | |
| RT | "Isolation of cDNA sequences coding for a part of human tissue | | | | |
| RT | plasminogen activator." | | | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 80:349-352 (1983). | | | | |
| RN | [8] | | | | |
| RP | SEQUENCE OF 1-36 FROM N.A. | | | | |
| RP | MEDLINE=85289338; PubMed=3161893; | | | | |
| RA | Fisher K., Waller E.K., Grosel G., Thompson D., Tizard R., | | | | |

RA Schleuning W.-D.;
 RT "Isolation and characterization of the human tissue-type plasminogen
 RT activator structural gene including its 5' flanking region.";
 RL J. Biol. Chem. 260:11223-11230(1985).
 RN [9]
 RP SEQUENCE OF 31-562 FROM N.A.
 RX MEDLINE=91291340; PubMed=1366681;
 RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;
 RT "Purification and characterization of tissue plasminogen activator
 RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells.";
 RL Agric. Biol. Chem. 55:1225-1232(1991).
 RN [10]
 RP SEQUENCE OF 36-562.
 RC TISSUE=Melanoma;
 RX MEDLINE=85000468; PubMed=6433976;
 RA Pohl G., Kaelin-Lang A., Bergsdorf N., Wallen P., Joernvall H.;
 RT "Tissue plasminogen activator: peptide analyses confirm an indirectly
 RT derived amino acid sequence, identify the active site serine residue,
 RT establish glycosylation sites, and localize variant differences.";
 RL Biochemistry 23:3701-3707(1984).
 RN [11]
 RP SEQUENCE OF 33-52 AND 311-330.
 RC TISSUE=Melanoma;
 RX MEDLINE=83209620; PubMed=6682760;
 RA Wallen P., Pohl G., Bergsdorf N., Raahy M., Ny T., Joernvall H.;
 RT "Purification and characterization of a melanoma cell plasminogen
 RT activator.";
 RL Eur. J. Biochem. 132:681-686(1983).
 RN [12]
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).
 RC TISSUE=umbilical vein;
 RX MEDLINE=90192129; PubMed=2107528;
 RA Siedert P.D., Fong K.;
 RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
 RT human endothelial cells.";
 RL Nucleic Acids Res. 18:1086-1086(1990).
 RN [13]
 RP STRUCTURE OF CARBOHYDRATES.
 RX MEDLINE=90092112; PubMed=2513186;
 RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;
 RT "Carbohydrate structure of recombinant human uterine tissue
 RT plasminogen activator expressed in mouse epithelial cells.";
 RL Eur. J. Biochem. 186:273-286(1989).
 RN [14]
 RP CARBOHYDRATE-LINKAGE SITE THR-96.
 RX MEDLINE=91159408; PubMed=1900431;
 RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;
 RT "Tissue plasminogen activator has an O-linked fucose attached to
 RT threonine-61 in the epidermal growth factor domain.";
 RL Biochemistry 30:2311-2314(1991).
 RN [15]
 RP DISULFIDE BONDS IN KRINGLE 2.
 RX MEDLINE=91244765; PubMed=1645336;
 RA Vlahos C.J., Wilhelm O.G., Hassell T., Jaskunas S.R., Bang N.U.;
 RT "Disulfide pairing of the recombinant kringle-2 domain of tissue
 RT plasminogen activator produced in *Scherichia coli*.";
 RL J. Biol. Chem. 266:10070-10072(1991).
 RN [16]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=96200985; PubMed=8613982;
 RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
 RT "The 2.3 Å crystal structure of the catalytic domain of recombinant
 RT two-chain human tissue-type plasminogen activator.";
 RL J. Mol. Biol. 258:117-135(1996).
 RN [17]
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=97449126; PubMed=9305622;
 RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,
 RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray
 RT crystal structure of single-chain human tPA.";
 RL EMBO J. 16:4797-4805(1997).

RN [18]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.
 RX MEDLINE=92118803; PubMed=1310033;
 RA de Vos A., Ullrich M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,
 RA Westbrook M.L., Kosiakof A.A.;
 RT "Crystal structure of the kringle 2 domain of tissue plasminogen
 RT activator at 2.4-Å resolution.";
 RL Biochemistry 31:270-279(1992).
 RN [19]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=90122799; PubMed=2558718;
 RA Byeon I.-J., Kelley R.F., Llinas M.;
 RT "1H NMR structural characterization of a recombinant kringle 2 domain
 RT from human tissue-type plasminogen activator.";
 RL Biochemistry 28:9350-9360(1989).
 RN [20]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=9120042; PubMed=1901789;
 RA Byeon I.-J., Kelley R.F., Llinas M.;
 RT "Kringle-2 domain of the tissue-type plasminogen activator. 1H-NMR
 RT assignments and secondary structure.";
 RL Eur. J. Biochem. 197:155-165(1991).
 RN [21]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=92106329; PubMed=1762144;
 RA Byeon I.-J., Llinas M.;
 RT "Solution structure of the tissue-type plasminogen activator kringle
 RT 2 domain complexed to 6-aminohexanoic acid an antifibrinolytic
 RT drug.";
 RL J. Mol. Biol. 222:1035-1051(1991).
 RN [22]
 RP STRUCTURE BY NMR OF 38-85.
 RX MEDLINE=92292163; PubMed=1602484;
 RA Downing A.K., Driscoll P.C., Harvey T.S., Dudgeon T.J., Smith B.O.,
 RA Bacon M., Campbell I.D.;
 RT "The solution structure and backbone dynamics of the fibronectin type
 RT I and epidermal growth factor-like pair of modules of tissue-type
 RT plasminogen activator.";
 RL J. Mol. Biol. 225:821-833(1992).
 RN [23]
 RP STRUCTURE BY NMR OF 36-126.
 RX MEDLINE=96077104; PubMed=7582899;
 RA Smith B.O., Downing A.K., Driscoll P.C., Dudgeon T.J., Campbell I.D.;
 RT "The solution structure and backbone dynamics of the fibronectin type
 RT I and epidermal growth factor-like pair of modules of tissue-type
 RT plasminogen activator.";
 RL Structure 3:823-833(1995).
 CC -I- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOMEN PLASMINOGEN
 CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
 CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
 CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
 CC MANY OTHER PHYSIOLOGICAL EVENTS.
 CC -I- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -I- SUBUNIT: BINDS TO FIBRIN WITH HIGH AFFINITY. THIS INTERACTION

Query Match 11.1%; Score 154.5; DB 1; Length 562;
 Best Local Similarity 39.6%; Pred. No. 1.1e-05;
 Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 25 CFMNGHLYVEDDTSPAPGRCLNWLDAQSLASAPVS-----GAGHSYCRPDE 75
 DB 127 CIEQGISYGTWSTVSAESGAECTNW--NSSALAQKPYSGRRPDALRLGLNHNVCRAIPDR 184
 QY 76 DPGPWCVYSGEAGVPEKRPCEDLRCPETTS 106
 DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213

RESULT 3
 ID TPA_RAT STANDARD; PRT; 559 AA.
 AC P19637;

DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (c-pa)
GN PLAT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89170114; PubMed=3148445;
RA NY T., Leonardson G., Heuvel A.J.W.;
RT "Cloning and characterization of a cDNA for rat tissue-type
RT plasminogen activator";
RL DNA 7:671-677(1988).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=90130448; PubMed=2105315;
RA Feng P., Ohlsson M., Ny T.;
RT "The structure of the TATA-less rat tissue-type plasminogen activator
RT gene. Specific sequence divergences in the promoter predict
RT differences in regulation of gene expression.";
RL J. Biol. Chem. 265:2022-2027(1990).
CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOMEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 BGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M23697; AAA41812.1; -
DR EMBL, M31197; AAA42261.1; -
DR EMBL, M31185; AAA42261.1; JOINED.
DR EMBL, M31186; AAA42261.1; JOINED.
DR EMBL, M31187; AAA42261.1; JOINED.
DR EMBL, M31188; AAA42261.1; JOINED.
DR EMBL, M31189; AAA42261.1; JOINED.
DR EMBL, M31190; AAA42261.1; JOINED.
DR EMBL, M31191; AAA42261.1; JOINED.
DR EMBL, M31192; AAA42261.1; JOINED.
DR EMBL, M31193; AAA42261.1; JOINED.
DR EMBL, M31194; AAA42261.1; JOINED.
DR EMBL, M31195; AAA42261.1; JOINED.
DR EMBL, M31196; AAA42261.1; JOINED.
DR EMBL, A19618; CAA01482.1; -
DR PIR, A31597; A31597.
DR HSPF, A35029; A35029.
DR HSPF, P00750; IRTF.
DR MEROPS, S01.232; -
DR InterPro, IPR001314; Chymotrypsin.

DR InterPro, IPR000561; EGF-like.
DR InterPro, IPR000083; Fibrinctn.
DR InterPro, IPR000001; Kringle.
DR InterPro, IPR001254; Ser_protease_Try.
DR Pfam, PF00008; EGF; 1.
DR Pfam, PF00039; fnl; 1.
DR Pfam, PF00051; kringle; 2.
DR Pfam, PF00089; trypsin; 1.
DR PRINTS, PR00722; CHYMOTRYPSIN.
DR PRINTS, PR00018; KRINGLE.
DR ProDom, PD000395; kringle; 2.
DR SMART, SM00181; EGF; 1.
DR SMART, SM00058; FN1; 1.
DR SMART, SM00130; KR; 2.
DR SMART, SM00020; Tryp_Spc; 1.
DR PROSITE, PS00022; EGF_1; 1.
DR PROSITE, PS01186; EGF_2; 1.
DR PROSITE, PS01253; FIBRONECTIN_1; 1.
DR PROSITE, PS00021; KRINGLE_1; 2.
DR PROSITE, PS00070; KRINGLE_2; 2.
DR PROSITE, PS00240; TRYPSIN_DOM; 1.
DR PROSITE, PS00134; TRYPSIN_HIS; 1.
DR PROSITE, PS00135; TRYPSIN_SER; 1.
KW plasminogen activator; Hydrolyase; Serine protease; Glycoprotein;
KW plasma; kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17
FT PROPEP 18 29
FT CHAIN 30 559
FT CHAIN 30 308
FT CHAIN 309 559
FT DOMAIN 36 78
FT DOMAIN 79 117
FT DOMAIN 124 205
FT DOMAIN 213 294
FT DOMAIN 309 559
FT ACT SITE 355 355
FT ACT SITE 404 404
FT ACT SITE 510 510
FT DISULFID 38 68
FT DISULFID 66 75
FT DISULFID 83 94
FT DISULFID 88 105
FT DISULFID 107 116
FT DISULFID 124 205
FT DISULFID 145 187
FT DISULFID 176 200
FT DISULFID 213 294
FT DISULFID 234 276
FT DISULFID 265 289
FT DISULFID 297 428
FT DISULFID 340 356
FT DISULFID 348 417
FT DISULFID 442 516
FT DISULFID 474 490
FT DISULFID 506 534
FT CARBOHYD 149 149
FT CARBOHYD 481 481
FT CONFLICT 380 380
SQ SEQUENCE 559 AA; 62903 MW; 7DBD1809CD1C21 CRG64;
Query Match 11.1%; Score 154; DB 1; Length 559;
Best Local Similarity 32.4%; Pred. No. 1, 2e-05;
Matches 45; Conservative 14; Mismatches 54; Indels 26; Gaps 6;
OY 25 CPMDNGLYREDOTSPAPRLCLMLDQSGLAGAPV-----GAGNHSYCNPPDE 75
DB 124 CFEGGQITTRGWTSTAKNGACILN--NSSALSGKPSASRPAKILGLGNHNYCNPPR 181
OY 76 DPRGMCYVSGAGVPEKPCEDLRCPE-----TTSQALPAFTTEIQEASEGFG 124
DB 182 DVR-PWCYVF-KAGKYTEFSGTPACPKGPTEDCYGKGVYRGTHFTT--SKASCLPW 237

QY 125 ADEVQVFAFAPALPARSEA 143
DB 238 NSMILIGKTYTAMRANSOA 256

RESULT 4
HGFA_MOUSE
ID HGFA_MOUSE STANDARD; PRT; 653 AA.
AC Q9R098; Q90KV4; 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (BC 3.4.21.-) (HGF activator) (HGFA).
GN HGFA_MOUSE
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Itosh H., Kataoka H., Koono H.;
RT "Mouse hepatocyte growth factor activator."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RA van Adelsberg J.S., Sehgal S., Kukes A., Brady C., Barrach J., Yang J., Han Y.;
RT "Activation of HGF by endogenous HGF activator is required for metanephric kidney morphogenesis in vitro."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM (BY SIMILARITY).
CC -1- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by, and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF099017; AAF02489.1; -
CC EMBL; AF224724; AAF34712.1; -
CC HSSP; P00763; IDPO.
CC MEROPS; S01.228; -
CC MGD; MGI:1859281; Hgfac.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR000562; FN Type II.
DR InterPro; IPR000083; Fibrinctn.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTRYPR1.

DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000955; FN Type_II; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydroxylase; glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 29
FT PROPEP 30 369
FT CHAIN 370 405
FT CHAIN 406 653
FT CHAIN 653 145
FT DOMAIN 157 195
FT DOMAIN 197 237
FT DOMAIN 238 276
FT DOMAIN 283 364
FT DOMAIN 406 653
FT ACT SITE 445 445
FT ACT SITE 496 495
FT ACT SITE 595 596
FT DISULFID 105 130
FT DISULFID 119 145
FT DISULFID 161 172
FT DISULFID 166 183
FT DISULFID 185 194
FT DISULFID 199 227
FT DISULFID 225 234
FT DISULFID 242 253
FT DISULFID 247 264
FT DISULFID 266 275
FT DISULFID 283 364
FT DISULFID 304 346
FT DISULFID 335 359
FT DISULFID 392 519
FT DISULFID 430 446
FT DISULFID 438 508
FT DISULFID 533 602
FT DISULFID 565 581
FT DISULFID 592 620
FT CARBOHYD 39 39
FT CARBOHYD 47 47
FT CARBOHYD 63 63
FT CARBOHYD 287 287
FT CARBOHYD 466 466
FT CARBOHYD 544 544
FT CONFLICT 164 164
SQ SEQUENCE 653 AA; 70567 MW; 88B4B2025DF7DC CRC64;
Query Match 10.8%; Score 150; DB 1; Length 653;
Best Local Similarity 32.6%; Pred. No. 3; le-05;
Matches 46; Conservative 17; Mismatches 46; Indels 32; Gaps 7;
QY 25 CPWMDGHLRYEDQTSAPAGLCLNW-----LDAQSLASAPVSGAGNHSYCRNPDED 76
DB 283 CFLGNGTEYRGVASTAAGLSCLAMNSDLXYOEJHVDG-VAAAVLLGLGPHAYCRPDDX 341
QY 77 PRGPGVCVSGEAGVEKRPCCDLRCPEFTTSOALPAFTTEIQEASBEGGADDEVQVFAFAPANA 136
DB 342 ER-FWCIYVVKDIALSWE-----YCRLTACESLARVHSQTEP-----ILA--A 380

OY 137 LPARSEAAVQPVIGISQRR 157
 DB 381 LP--ESAPAVRPTGKRRKR 399

RESULT 5
 ID TPA_MOUSE STANDARD; PRT; 559 AA.
 AC p1234;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (CPA)
 DE (c-PA) (c-plasminogen activator).
 GN PLAT.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8087303; PubMed=2826484;
 RA Rickles R.J., Darow A.L., Strickland S.;
 RT "Molecular cloning of complementary DNA to mouse tissue plasminogen
 RT activator mRNA and its expression during F9 teratocarcinoma cell
 RT differentiation.";
 RL J. Biol. Chem. 263:1563-1569(1988).
 CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOMEN PLASMINOGEN
 CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
 CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
 CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
 CC MANY OTHER PHYSIOLOGICAL EVENTS.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
 CC BOND.
 CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
 CC -1- PFM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
 CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
 CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
 CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
 CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE 1 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi-sb.ch/announce/>
 CC or send an email to license@ebi-sb.ch).
 CC -----
 DR EMBL; J03520; AAA40470.1; --
 DR PIR; A29941; A29941.
 DR HSSP; P00750; 1A5H.
 DR MEROPS; S01.232; --
 DR MGD; MGI:97640; Plac.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000083; Fibrinchnl.
 DR InterPro; IPR000001; Fibrinchnl.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF_1.
 DR Pfam; PF00039; fnl_1.
 DR Pfam; PF00051; kringle_2.
 DR Pfam; PF00089; trypsin_1.
 DR PRINTS; PR00722; CHYMOTRYPsin.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 2.

DR SMART; SM00181; EGF_1.
 DR SMART; SM00058; FNI; 1.
 DR SMART; SM00130; KR_2.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00021; KRINGLE_1; 2.
 DR PROSITE; PS50070; KRINGLE_2; 2.
 DR PROSITE; PS50070; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 DR PLASMINOGEN activation; Hydrolyse; Serine protease; Glycoprotein;
 KM Plasminogen activation; Hydrolyse; Serine protease; Glycoprotein;
 FT SIGNAL; 1 17
 FT PROPEP 18 29
 FT CHAIN 30 559
 FT CHAIN 30 308
 FT CHAIN 309 559
 FT DOMAIN 36 78
 FT DOMAIN 79 117
 FT DOMAIN 124 205
 FT DOMAIN 213 294
 FT DOMAIN 309 559
 FT ACT_SITE 355 355
 FT ACT_SITE 404 404
 FT ACT_SITE 510 510
 FT DISULFID 38 68
 FT DISULFID 66 75
 FT DISULFID 83 94
 FT DISULFID 88 105
 FT DISULFID 107 116
 FT DISULFID 124 205
 FT DISULFID 145 187
 FT DISULFID 176 200
 FT DISULFID 213 294
 FT DISULFID 234 276
 FT DISULFID 265 289
 FT DISULFID 297 428
 FT DISULFID 340 356
 FT DISULFID 348 417
 FT DISULFID 442 516
 FT DISULFID 474 490
 FT DISULFID 506 534
 FT CARBOHYD 149 149
 FT CARBOHYD 481 481
 SQ SEQUENCE 559 AA; 63110 MW; 4ACE57DC6A282A5 CRC64;

Query Match 10.6%; Score 147.5; DB 1; Length 559;
 Best Local Similarity 37.0%; Pred. No. 4,3e-05;
 Matches 34; Conservative 11; Mismatches 34; Indels 13; Gaps 4;

OY 25 CFWDNGHYREDQSPARGRCINWLDAGSGASAPVS-----GAGNHSYCRNDE 75
 DB 124 CFEEGDTTRGTSTWASGAECLNW--NSSVLSLKYNRRPAIKLGHNHNYCRNDR 181

OY 76 DRRGFWCVSGEAGVPEKRCEDLRCPETTSQ 107
 DB 182 DLK-FWCVGF-KAGKYTFTEFCSTPACPKKSE 211

RESULT 6
 HGFA HUMAN STANDARD; PRT; 655 AA.
 ID HGFA_HUMAN Q04756; Q14726;
 AC Q04756; Q14726;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGFA)
 DE activator (HGFA).
 GN HGFA.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 RX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Liver, and Serum.
 RX MEDLINE=93252878; PubMed=763665;
 RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,
 RA Kitamura N.;
 RT "Molecular cloning and sequence analysis of the cDNA for a human
 RT serine protease responsible for activation of hepatocyte growth
 RT factor: structural similarity of the protease precursor to blood
 RT coagulation factor XII.";
 RL J. Biol. Chem. 268:10024-10028(1993).
 RN [2]
 RP SEQUENCE OF 40-655 FROM N.A.
 RA Zhao S., Odeh C.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY
 CC CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM.
 CC -1- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
 CC DISULFIDE BOND.
 CC -1- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
 CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM.
 CC -1- TISSUE SPECIFICITY: LIVER.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isdb-sib.ch/announce/>
 CC or send an email to license@sib-sb.ch).
 CC -----
 CC EMBL: D14012; BAO3113.1; -
 CC EMBL: Z69923; CA93803.1; -
 CC F1R; A46688; A46688.
 CC HSSP: P00763; IDPO.
 CC MEROPS: S01.228; -
 CC Genew; HGNC:4894; HGFAc.
 DR MIM; 604552; -
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000742; EGF-2.
 DR InterPro; IPR000562; FN_Type_II.
 DR InterPro; IPR000083; Fibrinctn.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00013; ENTPEPIT.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; kringle; 1.
 DR ProDom; PD000995; FN_Type_II; 1.
 DR SMART; SM00181; EGF_2; 1.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRYPSIN; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS0186; EGF_2; 2.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.

DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 KW Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
 KW EGF-like domain; Repeat; Zymogen.
 FT SIGNAL 1 30
 FT PROPEP 31 372
 FT CHAIN 373 407
 FT CHAIN 408 655
 FT DOMAIN 108 148
 FT DOMAIN 160 198
 FT DOMAIN 200 240
 FT DOMAIN 241 279
 FT DOMAIN 286 367
 FT ACT_SITE 408 655
 FT ACT_SITE 447 497
 FT ACT_SITE 497 497
 FT ACT_SITE 598 598
 FT DISULFID 108 133
 FT DISULFID 122 148
 FT DISULFID 164 175
 FT DISULFID 169 186
 FT DISULFID 188 197
 FT DISULFID 202 230
 FT DISULFID 228 237
 FT DISULFID 245 256
 FT DISULFID 250 267
 FT DISULFID 269 278
 FT DISULFID 286 367
 FT DISULFID 307 349
 FT DISULFID 338 362
 FT DISULFID 394 521
 FT DISULFID 432 448
 FT DISULFID 440 510
 FT DISULFID 535 604
 FT DISULFID 567 583
 FT DISULFID 594 622
 FT CARBOHYD 48 48
 FT CARBOHYD 230 290
 FT CARBOHYD 468 468
 FT CARBOHYD 492 492
 FT CARBOHYD 546 546
 FT CONFLICT 644 644
 SQ SEQUENCE 655 AA; 70681 MW; 2CF2FE1B862ED7 CRC64;
 Query Match 10.5%; Score 146.5; DB 1; Length 655;
 Best local Similarity 36.9%; Pred.No.6.2e-05;
 Matches 41; Conservative 10; Mismatches 41; Indels 19; Gaps 6;
 QY 25 CFWDNGLHREDSPTSPAGRLCLN-----LDAOSGLASPVSGAGNSYCRNPED 76
 DB 286 CPLNGTGRGVASTASGLSLANWSDLYQELHVDV-VGAALALGLCPHAYCNPND 344
 QY 77 PRGPWCYVSGEAGVP-----EKRPCEDLRCPETSOALPATTICE-ASEG 122
 DB 345 ER-PWCYVVKDSALSMEXCRLEACESL-----TRVQLSPDLLATLPAPAPG 390
 RESULT 7
 P1MN HUMAN STANDARD; PRT; 810 AA.
 AC P00747;
 DT 21-JUN-1986 (Rel. 01, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Plasmaogen precursor (BC 3.4.21.7) [Contains: Angiostatin].
 GN PEG.
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90202879; PubMed=2318848;
RA Peteren T.E., Martzen M.R., Ichinose A., Davie E.W.;
RT "Characterization of the gene for human plasminogen, a key proenzyme
RL in the fibrinolytic system.";
RL J Biol. Chem. 265:6104-6111(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97162490; PubMed=3030813;
RA Forstgen M., Raden B., Israelsen M., Larsen K., Heden L.-O.;
RT "Molecular cloning and characterization of a full-length cDNA clone
RL for human plasminogen.";
RL FEBS Lett. 213:254-260(1987).
RN [3]
RP SEQUENCE OF 20-810.
RA Sottcup-Jensen L., Petersen T.E., Magnusson S.;
RL Submitted (JUL-1977) to the PIR data bank.
RN [4]
RP SEQUENCE OF 292-810 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
RL human and bovine plasminogen.";
RL Biochemistry 23:4243-4250(1984).
RN [5]
RP SEQUENCE OF 20-100.
RX MEDLINE=75093329; PubMed=122932;
RA Wiman B., Wallen P.;
RT "Structural relationship between 'glutamic acid' and 'lysine' forms
RT of human plasminogen and their interaction with the NH2-terminal
RL activation peptide as studied by affinity chromatography.";
RL Eur. J. Biochem. 50:489-494(1975).
RN [6]
RP SEQUENCE OF 95-580; 581-626; 657-700 AND 722-810.
RA Sottcup-Jensen L., Claes H., Zajdel M., Petersen T.E., Magnusson S.;
RT (In) Davidson J.F., Rowan R.M., Samama M.M., Desnoyers P.C. (eds.);
RL progress in chemical fibrinolysis and thrombolysis, Pp.3:191-209,
RL Raven Press, New York (1978).
RN [7]
RP SEQUENCE OF 483-604.
RX MEDLINE=76043692; PubMed=126863;
RA Wiman B., Wallen P.;
RT "Amino-acid sequence of the cyanogen-bromide fragment from human
RL plasminogen that forms the linkage between the plasmin chains.";
RL Eur. J. Biochem. 58:539-547(1975).
RN [8]
RP SEQUENCE OF 581-810.
RX MEDLINE=77225245; PubMed=142009;
RA Wiman B.;
RT "Primary structure of the B-chain of human plasmin.";
RL Eur. J. Biochem. 76:129-137(1977).
RN [9]
RP ACTIVE SITE.
RX MEDLINE=73149248; PubMed=4694729;
RA Robbins K.C., Bernabe P., Arzadon L., Summaria L.;
RT "The primary structure of human plasminogen. II. The histidine loop
RL of human plasmin: light (b) chain active center histidine sequence.";
RL J Biol. Chem. 248:1631-1633(1973).
RN [10]
RP ACTIVE SITE.
RX MEDLINE=69234739; PubMed=4240117;
RA Grotskopf W.R., Summaria L., Robbins K.C.;
RT "Studies on the active center of human plasmin. Partial amino acid
RL sequence of a peptide containing the active center serine residue.";
RL J Biol. Chem. 244:3590-3597(1969).
RN [11]
RP OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.
RX MEDLINE=82213905; PubMed=6919539;
RL Trexler M., Vail Z., Pecthy L.;

RT "Structure of the omega-aminocarboxylic acid-binding sites of human
RT plasminogen. Arginine 70 and aspartic acid 56 are essential for
RT binding of ligand by kringle 4.";
RL J Biol. Chem. 257:7401-7406(1982).
RN [12]
RP FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.
RX MEDLINE=85054784; PubMed=6094526;
RA Vail Z., Pecthy L.;
RT "The fibrin-binding site of human plasminogen. Arginines 32 and 34
RL are essential for fibrin affinity of the kringle 1 domain.";
RL J Biol. Chem. 259:13690-13694(1984).
RN [13]
RP PHOSPHORYLATION SITE SER-597.
RX MEDLINE=97345939; PubMed=9201958;
RA Wang H., Proctor M., Bretthauer R.K., Castellino F.J.;
RT "Serine-578 is a major phosphorylation locus in human plasma
RL plasminogen.";
RL Biochemistry 36:8100-8106(1997).
RN [14]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=88185329; PubMed=3356193;
RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamberling J.P.,
RA Gerwig G.J., van Halbeek H., Vliegenhart J.F.;
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
RT plasminogen. Species specificity in relation to alkylation and
RL fucosylation patterns.";
RL Eur. J. Biochem. 173:57-63(1988).
RN [15]
RP CARBOHYDRATE-LINKAGE SITE SER-268.
RX MEDLINE=97207306; PubMed=9054441;
RA Pizzo-Shepherd S.R., Stevens R.D., Andon N.L., Enghild J.J.,
RA Pizzo S.V.;
RT "Evidence for a novel O-linked sialylated triasaccharide on Ser-248 of
RL human plasminogen 2.";
RL J Biol. Chem. 272:7408-7411(1997).
RN [16]
RP CHARACTERIZATION OF ANGIOSTATIN AND PARTIAL SEQUENCE.
RX MEDLINE=95042728; PubMed=7525077;
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA Moses H., Lane W.S., Cao Y., Sage E.H., Folkman R.;
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
RL suppression of metastases by a Lewis lung carcinoma.";
RL Cell 79:315-328(1994).
RN [17]
RP CHARACTERIZATION OF ANGIOSTATIN.
RX MEDLINE=97238710; PubMed=9102221;
RA Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Maden J.W.,
RA Lapevych R., Nacy C.A.;
RT "A recombinant human angiostatin protein inhibits experimental primary
RL and metastatic cancer.";
RL Cancer Res. 57:1329-1334(1997).
RN [18]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.
RX MEDLINE=92031502; PubMed=1657148;
RA Mulichak A.M., Tulinsky A., Ravichandran K.G.;
RT "Crystal and molecular structure of human plasminogen kringle 4
RL refined at 1.9-A resolution.";
RL Biochemistry 30:10576-10588(1991).
RN [19]
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.
RX MEDLINE=92031503; PubMed=1657149;
RA Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;
RT "The refined structure of the epsilon-aminocaproic acid complex of
RL human plasminogen kringle 4.";
RL Biochemistry 30:10589-10594(1991).
RN [20]
RP X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
RX Stee B., Yamano A., Whitlow M., Teeter W.M.;
RT "Structure of human plasminogen kringle 4 at 1.68 Angstrom and 277 K.
RL Acta Crystallogr. D 53:169-178(1997).
RN [21]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.


```

RX MEDLINE-96180681; PubMed=8611560;
RA Mathews J.I., Vanderhoff-Hanover P., Castellino F.J., Tullinsky A.;
RT "Crystal structures of the recombinant kringle 1 domain of human
RT plasminogen in complexes with the ligands epsilon-aminocaproic acid
RT and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";
RL Biochemistry 35:2567-2576(1996).
RP X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.
RX MEDLINE-96198034; PubMed=9521645;
RA Chang J., Mochalkin I., McNamee S.G., Cheng B., Tullinsky A.,
RA Castellino F.J.;
RT "Structure and ligand binding determinants of the recombinant kringle
RT 5 domain of human plasminogen.";
RL Biochemistry 37:3258-3271(1998).
RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE-94237157; PubMed=8181475;
RA Rejzante M.R., Llinas M.;
RT "1H-NMR assignments and secondary structure of human plasminogen
RT kringle 1.";
RL Eur. J. Biochem. 221:927-937(1994).
RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE-94237158; PubMed=8181476;
RA Rejzante M.R., Llinas M.;
RT "Solution structure of the epsilon-aminohexanoic acid complex of
RT human plasminogen kringle 1.";
RL Eur. J. Biochem. 221:939-949(1994).
RP STRUCTURE BY NMR OF 183-354.
RX MEDLINE-96194156; PubMed=8652577;
RA Soehndel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,
RA Rickli E.B.;
RT "Recombinant gene expression and 1H NMR characteristics of the
RT kringle (2 + 3) supermodule: spectroscopic/functional individuality
RT of plasminogen kringle domains";
RL Biochemistry 35:2357-2364(1996).
RP STRUCTURE BY NMR OF 374-461.
RX MEDLINE-90219023; PubMed=2157850;
RA Atkinson R.A., Williams R.J.P.;
RT "Solution structure of the kringle 4 domain from human plasminogen by
RT 1H nuclear magnetic resonance spectroscopy and distance geometry.";
RL J. Mol. Biol. 212:541-552(1990).
RP VARIANTS PHE-374 AND THR-620.
Query Match 10.2%; Score 142; DB 1; Length 810;
Best Local Similarity 37.2%; Pred. No. 0.00019;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;
QY 25 CPWDNGHLYREDPOTSPAPGRLCINMLDA-----QSGLASAPVSGAGNHSYCNPPEDPRG 79
DB 103 CKTGNGKNGYRGTMSKTKNGJTCOKMSSTSPHRFRSPATHPSEG-L-ENNYCNPNDPOG 161
QY 80 PWCYVGEAGVPERK--PCEDLRCPCE 103
DB 162 PWCYTTD-----PEKRYDYCDILECEE 183
RESULT 8
PLMN RAT STANDARD; PRT; 169 AA.
AC 001177;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasminogen (EC 3.4.21.7) (Fragment).
GN PLG
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE-91250378; PubMed=1645711;
RA Kenalas J.J., Møller S.P.;
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT receptor site for plasminogen.";
RL J. Biol. Chem. 266:10825-10829(1991).
CC -1- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION
CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UKOKINASE-TYPE PLASMINOGEN
CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT PROTEINS, SUCH
CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC LAMININ AND VON WILLEBRAND FACTOR.
CC -1- CATALYTIC ACTIVITY: preferential cleavage: Lys-[Xaa > Arg]-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -1- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC -1- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; M62832; AAA41884.1; --
CC PIR; A40522; A40522.
CC HSRP; P00747; 1PMK.
CC MEMOS; S01.233; --
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser. protease_Try.
CC Pfam; PF00051; Kringle_2.
CC Pfam; PF00095; Kringle_2.
CC DR SMART; SM00130; KR; 1.
CC DR PROSITE; PS00021; KRINGLE_1; 1.
CC DR PROSITE; PS0070; KRINGLE_2; 2.
CC DR PROSITE; PS0070; TRYPSIN_DOM; PARTIAL.
CC DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
CC DR PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
CC KW Hydrolyase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
CC Tissue remodeling; Blood coagulation; Kringle; Repeat.
CC FT NON_TER 1 1
CC FT DOMAIN 1 10 KRINGLE 3 (BY SIMILARITY).
CC FT DOMAIN 34 112 KRINGLE 4 (BY SIMILARITY).
CC FT DOMAIN 139 >169 KRINGLE 5 (BY SIMILARITY).
CC FT DISULFID 34 112 BY SIMILARITY.
CC FT DISULFID 55 95 BY SIMILARITY.
CC FT DISULFID 83 107 BY SIMILARITY.
CC FT NON_TER 169 169
SQ SEQUENCE 169 AA; 18401 MW; 77A54214C49D010C CRC64;
Query Match 9.9%; Score 138; DB 1; Length 169;
Best Local Similarity 31.8%; Pred. No. 7.2e-05;
Matches 35; Conservative 13; Mismatches 48; Indels 14; Gaps 5;
QY 25 CPWDNGHLYREDPOTSPAPGRLCINMLDA-----QSGLASAPVSGAGNHSYCNPPEDPRG 79
DB 34 CYGNGKSYRGRSTNTNGKKGQSWMTSPHSHKTPANFPDGL-ENNYCNPNDPOG 92
QY 80 PWCYVGEAGVPERK--PCEDLRCPCE 103
DB 93 PWCYTTD-----PSVMEYCNLAKCSSTGGV--ASALVYQVVSARCTSE 136

```



```

DR MEROPS, S01.231; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser.protease_Try.
DR Pfam: PF00051; kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR ProDom: PD000395; Kringle; 1.
DR SMART: SM00130; KR_1.
DR SMART: SM00020; TRYP_SPC; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; FALSE_NEG.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS50070; KRINGLE_2; 1.
DR PROSITE: PS50240; TRYPsin DOM; 1.
DR PROSITE: PS00134; TRYPsin HIS; 1.
DR PROSITE: PS00135; TRYPsin SER; 1.
DR Plasmaprotease activation: Hydrolyase; Serine protease; Glycoprotein;
KR Kringle; EGF-like domain; Zymogen; signal.
KM SIGNAL 1 20
FT CHAIN 1 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 188 CHAIN A (BY SIMILARITY).
FT CHAIN 190 442 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 189 CONNECTING PEPTIDE.
FT DOMAIN 190 442 SERINE PROTEASE.
FT CATHID 152 152 N-LINKED (GLCMAC. . .).
FT CATHID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 179 220 INTERCHAIN (BY SIMILARITY).
FT DISULFID 220 236 BY SIMILARITY.
FT DISULFID 228 289 BY SIMILARITY.
FT DISULFID 324 393 BY SIMILARITY.
FT DISULFID 356 372 BY SIMILARITY.
FT DISULFID 383 411 BY SIMILARITY.
FT ACT_SITE 235 235 CHARGE RELAY SYSTEM.
FT ACT_SITE 286 286 CHARGE RELAY SYSTEM.
FT ACT_SITE 387 387 CHARGE RELAY SYSTEM.
FT ACT_SITE 241 241 Q -> H (IN REF. 1; CAA25806).
FT CONFLICT 242 242 Q -> H (IN REF. 1; CAA25806).
FT CONFLICT 288 288 A -> GS (IN REF. 1; CAA25806).
SQ SEQUENCE 442 AA; 4916 MW; E832CFE501321EE CRC64;

Query Match 9.7%; Score 135.5; DB 1; Length 442;
Beet Local Similarity 36.9%; Pred. No. 0.00035;
Matches 38; Conservative 7; Mismatches 37; Indels 21; Gaps 5;

OY 25 CWDNGHLYREDQTSPPAGLRCLINWDAQSGL---ASAPVS---GAGNHSYGRNFDDEP 77
DB 72 CFEQGHSHYRGKANTNTGGRPLCPMNSATVLTNTYHARPDALQLGKKNYCRNDP-NQ 130
OY 78 RGPWCYVS-----GEAGVP-----EKRPEDIRCEPETSQ 107
DB 131 RRPWCYVGVGLKQVQECWVPNCSGGSEHRRPVPYDKGNFSPTE 173

RESULT 11
URTB_DESRO STANDARD; PRT; 431 AA.
ID URTB_DESRO
AC P98121;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSBA
DE beta).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
NCBI_Textid=9430;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kaezschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=9393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kaezschmar J., Haendler B., Langer G., Baldus B., Wilt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC AGENT.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: MONOMER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@1sb-sib.ch).
CC -----
DR EMBL: M63989; AAA31594.1; -.
DR HSP: P98119; IAS1.
DR MEROPS; S01.239; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser.protease_Try.
DR Pfam: PF00089; EGF; 1.
DR Pfam: PF00051; kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 1.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; TRYP_SPC; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS50070; KRINGLE_2; 1.
DR PROSITE: PS50240; TRYPsin DOM; 1.
DR PROSITE: PS00134; TRYPsin HIS; 1.
DR PROSITE: PS00135; TRYPsin SER; 1.
DR Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
KM Kringle; EGF-like domain; signal; Multigene family.
KM SIGNAL 1 36
FT CHAIN 37 431 SALIVARY PLASMINOGEN ACTIVATOR BETA.
FT DOMAIN 37 163 EGF-LIKE.
FT DOMAIN 179 431 KRINGLE.
FT ACT_SITE 226 226 SERINE PROTEASE.
FT ACT_SITE 275 275 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 382 382 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 41 52 BY SIMILARITY.
FT DISULFID 46 63 BY SIMILARITY.
FT DISULFID 65 74 BY SIMILARITY.
FT DISULFID 82 163 BY SIMILARITY.

```


FT CONFLICT 403 403 N -> K (IN REF. 2).
 FT CONFLICT 417 417 Y -> H (IN REF. 2).
 FT CONFLICT 435 435 M -> R (IN REF. 2).
 SQ SEQUENCE 477 AA; 53719 MW; 17486555COE5077C CRC64;
 Query Match 9.6%; Score 134; DB 1; Length 477;
 Best Local Similarity 38.6%; Pred. No. 0.0005;
 Matches 27; Conservative 9; Mismatches 20; Indels 14; Gaps 3;
 QY 25 CFWNDGHLVREDQTSFAPGLCLNWLDAOSGL-----ASAPVSGAGHSTCRPPD 74
 DB 128 CYKDDGVTVRGWSTSSGACQINW---NSNLLTRRTYNGRRSDATITGLGHNYNCRPMD 184
 QY 75 EDPGPGWCYV 84
 DB 185 NNSK-FWCYV 193
 RESULT 13
 HGFL_MOUSE STANDARD; PRT; 716 AA.
 AC P26928;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hepatocyte growth factor-like protein precursor (Macrophage
 stimulatory protein) (MSP).
 GN MST1 OR HGFL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Liver;
 RX MEDLINE=92002017; PubMed=1832957;
 RA Fritzenr Degen S.J., Stuart L.A., Han S., Jamison C.S.;
 RT "Characterization of the mouse cDNA and gene coding for a hepatocyte
 growth factor-like protein: expression during development.";
 RL Biochem J 30:9781-9791(1991).
 CC -1- FUNCTION: PROBABLY HAS NO PROTEOLYTIC ACTIVITY, SINCE CRUCIAL AA
 CHARACTERISTIC OF SERINE PROTEASES CATALYTIC SITES ARE NOT
 CONSERVED.
 CC -1- TISSUE SPECIFICITY: LIVER. LOWER LEVELS IN LUNG, PLACENTA AND
 ADRENAL.
 CC -1- DEVELOPMENTAL STAGE: IS EXPRESSED AT LOW LEVELS DURING GESTATION.
 CC JUST BEFORE BIRTH THE LEVEL INCREASES DRAMATICALLY AND REMAINS
 STABLE AFTERWARDS.
 CC -1- PTM: MAY BE CLEAVED AFTER AA 488, TO YIELD A TWO-CHAIN MOLECULE
 HELD TOGETHER BY DISULFIDE BONDS, OR CLEAVED INTO TWO SEPARATE
 POLYPEPTIDES.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M74180; AAA50166.1; -;
 DR EMBL: M74181; AAA50167.1; -;
 DR HSSP: P00747; IREN.
 DR MEROPS: S01.975; -;
 DR MGD: MGI:96080; HGFL.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR003014; PAN.
 DR InterPro: IPR003609; Pan app.
 DR InterPro: IPR001254; Ser_Protease_Try.
 DR Pfam: PF00024; PAN; 1.

DR Pfam: PF00051; kringle; 4.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR Prodom: PD000395; Kringle; 4.
 DR SMART: SM00130; KR; 4.
 DR SMART: SM00473; PAN AP; 1.
 DR SMART: SM00020; Tryp_Spc; 1.
 DR PROSITE: PSS00021; KRINGLE 1; 4.
 DR PROSITE: PSS0070; KRINGLE 2; 4.
 DR PROSITE: PSS0240; TRYPSIN_DOM; 1.
 DR KRingle; Glycoprotein; Serine protease homolog; Repeat; Signal.
 KW KRingle; Glycoprotein; Serine protease homolog; Repeat; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 716
 FT DOMAIN 19 109
 FT DOMAIN 110 186
 FT DOMAIN 191 268
 FT DOMAIN 292 370
 FT DOMAIN 379 457
 FT DOMAIN 489 716
 FT DISULFID 56 78
 FT DISULFID 60 66
 FT DISULFID 110 186
 FT DISULFID 131 169
 FT DISULFID 157 181
 FT DISULFID 191 268
 FT DISULFID 194 333
 FT DISULFID 212 251
 FT DISULFID 240 263
 FT DISULFID 292 370
 FT DISULFID 313 352
 FT DISULFID 341 364
 FT DISULFID 379 457
 FT DISULFID 400 440
 FT DISULFID 428 452
 FT DISULFID 477 593
 FT DISULFID 512 528
 FT DISULFID 607 672
 FT DISULFID 637 651
 FT DISULFID 662 690
 FT CARBOHYD 72 72
 FT CARBOHYD 173 173
 FT CARBOHYD 305 305
 FT CARBOHYD 620 620
 FT CONFLICT 19 19
 SQ SEQUENCE 716 AA; 80598 MW; BDC02EF85213ACC CRC64;
 Query Match 9.6%; Score 134; DB 1; Length 716;
 Best Local Similarity 28.2%; Pred. No. 0.0008;
 Matches 42; Conservative 9; Mismatches 46; Indels 52; Gaps 7;
 QY 2 LLAWQAFVLSNMLAAVYSG-----GCFWNDGHLVREDQTSFAPGLCLNWL 49
 DB 60 LLPTQ-----HSLHTQLYHSSLSCHLFQKDYVTRCINDNGVSYRGYARNGSLPQAW 134
 QY 50 ---LDAOSGLASAPVSGAGHSTCRPPDPRGWCYVS----- 85
 DB 135 SRFPNDHKYTPYRKGL-BENFCRPDDBPGWPCYTTNSVRFQSGIKTCREAVCVL 193
 QY 86 -----GEAGVPEK-RPCE--DLRCPET 104
 DB 194 CNGEDYRGEVDVTFESGRECOAMDQHPHS 222
 RESULT 14
 URUG_DESRO STANDARD; PRT; 394 AA.
 AC P49150;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSFA
 gamma).

OS Desmodus rotundus (Vampire bat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 OC Desmodontinae; Desmodus.
 OC NCBI_TaxID=9430;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=92039036; PubMed=1937019;
 RA Kraeuschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
 RA Alagon A., Donner P., Schleuning W.D.;
 RT "The plasminogen activator family from the salivary gland of the
 RT vampire bat Desmodus rotundus: cloning and expression.";
 RL Gene 105:1229-237 (1991).
 [2]
 RP CHARACTERIZATION.
 RX MEDLINE=9393059; PubMed=1309059;
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
 RA Kraeuschmar J., Haendler B., Langer G., Baldus B., Wilt W.,
 RA Donner P.;
 RT "Plasminogen activators from the saliva of Desmodus rotundus (common
 RT vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 667:395-403 (1992)
 CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
 CC EXCLUSIVELY HAEMATOPOAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
 CC AGENT.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M63990; AAA1595.1; -;
 DR HSSP; P98119; IASI.
 DR MEROPS; S01.239; -;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRY_Spec; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KM Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 KM Kringle; Signal; Multigene family.
 FT SIGNAL 1 36
 FT CHAIN 1 36
 FT DOMAIN 37 394
 FT DOMAIN 142 126
 FT ACT_SITE 142 394
 FT ACT_SITE 189 189
 FT ACT_SITE 238 238
 FT ACT_SITE 345 345
 FT DISULFID 45 126
 FT DISULFID 66 108
 FT DISULFID 97 121
 FT DISULFID 131 262
 FT DISULFID 174 190
 FT DISULFID 182 251
 FT BY SIMILARITY.

FT DISULFID 276 351 BY SIMILARITY.
 FT DISULFID 308 324 BY SIMILARITY.
 FT DISULFID 341 369 BY SIMILARITY.
 FT CAROHYD 315 315 N-LINKED (GLCNAC...)(POTENTIAL).
 SQ SEQUENCE 394 AA; 44105 MW; 9CCD6F52FD81FCD CR664;
 Query Match 9.5%; Score 132; DB 1; Length 394;
 Best Local Similarity 32.7%; Pred. No. 0.0006;
 Matches 33; Conservative 9; Mismatches 39; Indels 20; Gaps 5;
 Oy 17 AEAYS--GGCFWDGCHYREDOTSPPAPGRCINW-----LDAQGLASAPVSGAGNH 67
 Db 35 SRAYGDPHATCYKQGGVYRGTMTSGAGCINMNSNLLIRTYNGRPEAVKGLGNH 94
 Oy 68 SYCRNPDEDPGPGWCVY-----SGEAYPEKRPEDLRC 101
 Db 95 NYCRRNPDGASK-PWCYVIKARKFTSECSVP---VCSKATC 131
 RESULT 15
 ID UROK_HUMAN STANDARD: PRT; 431 AA.
 AC P00729; Q15844; Q16618; Q969W6;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 20-MAR-1987 (Rel. 04, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
 DE (U-plasminogen activator).
 GN UPA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85215647; PubMed=2987867;
 RA Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blasi F.;
 RT "The human urokinase-plasminogen activator gene and its promoter";
 RT Nucleic Acids Res. 13:2759-2771 (1985).
 [2]
 RP SEQUENCE FROM N.A.
 RA Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,
 RA Steffens G.J., Haynaker H.L.;
 RT "Cloning and expression of the gene for pro-urokinase in Escherichia
 RT coli.";
 RL Biotechnology 3:923-929 (1985).
 [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86056954; PubMed=2415429;
 RA Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Arimura H.,
 RA Nishida M., Suyama T.;
 RT "Molecular cloning of cDNA coding for human preprourokinase.";
 RT Gene 36:183-188 (1985).
 [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85203359; PubMed=388571;
 RA Jacobs P., Cravador A., Lorian R., Brockly F., Colau B., Chuchana P.,
 RA van Elsen A., Herzog A., Bollen A.;
 RT "Molecular cloning, sequencing, and expression in Escherichia coli of
 RT human preprourokinase cDNA.";
 RL DNA 4:139-146 (1985).
 [5]
 RP SEQUENCE FROM N.A.
 RA Rieder M.-D., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
 RA Nickerson D.A.;
 RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 [6]
 RP SEQUENCE FROM N.A.
 RC Tissue:lung;
 RA Strauszberg R.;
 RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 [7]
 RP SEQUENCE OF 66-431 FROM N.A.

RX MEDLINE=84272706; PubMed=6589620;
 RA Verde P., Scoppelli M.P., Galeffi P., di Nocera P., Blasi F.;
 RT "Identification and primary sequence of an unspliced human urokinase
 RT poly(A)+RNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
 RN [8]
 RP SEQUENCE OF 21-177.
 RX MEDLINE=83055084; PubMed=6754569;
 RA Gunler W.A., Steffens G.J., Oetting F., Kim S.-M.A., Frankus E.,
 RA Flohe L.;
 RT "The primary structure of high molecular mass urokinase from human
 RT urine. The complete amino acid sequence of the A chain.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
 RN [9]
 RP SEQUENCE OF 156-176 AND 179-224.
 RX MEDLINE=83003608; PubMed=6749492;
 RA Schaller J., Nick H., Rickli E.E., Gillesen D., Lejgier W.,
 RA Studer R.O.;
 RT "Human low-molecular-weight urinary urokinase. Partial
 RT characterization and preliminary sequence data of the two polypeptide
 RT chains.";
 RL Eur. J. Biochem. 125:251-257(1982).
 RN [10]
 RP SEQUENCE OF 158-410.
 RX MEDLINE=83055099; PubMed=6754572;
 RA Steffens G.J., Gunzler W.A., Oetting F., Frankus E., Flohe L.;
 RT "The complete amino acid sequence of low molecular mass urokinase
 RT from human urine.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=96000858; PubMed=8591045;
 RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
 RA Dobson C.M., Stuart D.I., Jones E.Y.;
 RT "The crystal structure of the catalytic domain of human
 RT urokinase-type plasminogen activator.";
 RL Structure 3:681-691(1995).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
 RX MEDLINE=20266327; PubMed=10805774;
 RA Spel S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,
 RA Bode W., Maggolen V., Huber R., Moroder L.;
 RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly
 RT selective inhibitors of human urokinase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
 RN [13]
 RP STRUCTURE BY NMR.
 RX MEDLINE=89127526; PubMed=2536903;
 RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.;
 RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
 RT dimensional NMR.";
 RL Nature 337:579-582(1989).
 RN [14]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=93003110; PubMed=1327118;
 RA Li X., Smith R.A.G., Dobson C.M.;
 RT "Sequential 1H NMR assignments and secondary structure of the kringle
 RT domain from urokinase.";
 RL Biochemistry 31:9562-9571(1992).
 RN [15]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=94149701; PubMed=8107091;
 RA Li X., Bohman A.M., Llinas M., Smith R.A.G., Dobson C.M.;
 RT "Solution structure of the kringle domain from urokinase-type
 RT plasminogen activator.";
 RL J. Mol. Biol. 235:1548-1559(1994).
 RN [16]
 RP VARIANT LEU-141.
 RX MEDLINE=96186279; PubMed=8656531;
 RA Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,
 RA Sawasaki Y., Hanada K.;
 RT "Characterization of single chain urokinase-type plasminogen
 RT activator with a novel amino-acid substitution in the kringle

RT structure.";
 RL Biochim. Biophys. Acta 1293:83-89(1996).
 RN [17]
 RP VARIANT LEU-141.
 RX MEDLINE=97218551; PubMed=9065988;
 RA Come B., Berczy M., Belin D.;
 RT "Detection of polymorphisms in the human urokinase-type plasminogen
 RT activator gene.";
 RL Thromb. Haemost. 77:434-435(1997).
 RN [18]
 RP ERRATUM.
 RA Come B., Berczy M., Belin D.;
 RL Thromb. Haemost. 78:973-973(1997).
 RN [19]
 RP VARIANT LEU-141.
 RX MEDLINE=97379720; PubMed=9194591;
 RA Turkmen B., Schmitt M., Schmaljohr B., Trommler P., Hell W.,
 RA Crenzelburg S., Graeff H., Maggolen V.;
 RT "Mutational analysis of the gene encoding urokinase-type plasminogen
 RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
 RL Electrophoresis 18:686-689(1997).
 CC -1- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR
 CC THERAPY OF THROMBOLYTIC DISORDERS.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
 CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
 CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
 CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.
 CC -1- PHARMACEUTICAL: Available under the name Abbockinase (Abbot). Used
 CC in Pulmonary Embolism (PE) to initiate fibrinolysis.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; K02419; CA26268.1; -
 DR EMBL; M15476; AAA61252.1; -
 DR EMBL; D00244; BAA00175.1; -
 DR EMBL; D1143; BAA01915.1; -
 DR EMBL; X02760; CA26535.1; -
 DR EMBL; AF377330; AAK53822.1; -
 DR EMBL; BC033575; AAH13575.1; -
 DR EMBL; K03266; AAC97138.1; -
 DR EMBL; K02286; AAA61252.1; -
 DR EMBL; A21571; CAA01559.1; -
 DR EMBL; A18397; CAA01390.1; -
 DR PIR; A00931; UKHU.
 DR PIR; A32974; A32974.
 DR PDB; 1KDU; 31-OCT-93.
 DR PDB; 1LWV; 29-JAN-96.
 DR PDB; 1URK; 08-MAY-95.
 DR PDB; 1EIN; 17-MAY-00.
 DR MEROPS; S01.231; -
 DR GlycoSuiteDB; P00749; -
 DR Genew; HGNC:9052; PLAU.
 DR TrEMBL; IPR001314; Chymotrypsin.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; Cyspin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 1.

DR SMART: SM00130: KR: 1.
DR SMART: SM00020: TRY_SPC: 1.

Query Match 9.5%; Score 132; DB 1; Length 431;
Best Local Similarity 32.7%; Pred No 0.00067;
Matches 33; Conservative 13; Mismatches 33; Indels 22; Gaps 5;

QY 25 CFNDNGHLYREDOTSPABGLCLWLD-----QSGLASAPVSGAGNHSYCRNPDEDP 77
DB 70 CIEGNGHIFIKGASTUTWKRPELPMNSATVYLQOTYHHRSDPLQGLGKANYCRNPD-NR 128
QY 78 RGPWCYVSGEAGVPEKRP-----CEDLRCPETTSOAL 109
DB 129 RRPWCYV--QVGL--KPLVQECNVHDCA DGKKPSSPEEL 164

Search completed: April 7, 2003, 06:48:34
Job time : 12.8979 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 06:57:52 ; Search time 55.4987 Seconds
(without alignments)
976,427 Million cell updates/sec

Title: US-10-057-951-2
Perfect score: 1393
Sequence: 1 MLAWVOAFVSNMLAEAY.....PVDPOBSGTPMLGQAGTPGA 263

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Length | DB ID | Description |
|------------|-------|--------------|---------------|---------------------|
| 1 | 1393 | 100.0 | 263 4 Q96FE7 | Q96FE7 homo sapien |
| 2 | 1390 | 99.8 | 263 4 O00318 | O00318 homo sapien |
| 3 | 167 | 12.0 | 562 6 Q8SQ23 | Q8SQ23 sus scrofa |
| 4 | 154.5 | 11.1 | 516 4 Q9BU99 | Q9BU99 homo sapien |
| 5 | 150 | 10.8 | 653 11 Q8VCS4 | Q8VCS4 mus musculu |
| 6 | 147.5 | 10.6 | 559 11 Q8VCS4 | Q8VCS4 mus musculu |
| 7 | 142 | 10.2 | 810 4 Q15146 | Q15146 homo sapien |
| 8 | 139.5 | 10.0 | 385 5 Q25101 | Q25101 hardmania m |
| 9 | 137 | 9.8 | 716 11 P70521 | P70521 rattus norv |
| 10 | 136.5 | 9.8 | 420 13 Q90504 | Q90504 eptatretus |
| 11 | 135 | 9.7 | 112 11 Q90504 | Q90504 rattus norv |
| 12 | 134.5 | 9.7 | 395 4 Q9BZM1 | Q9BZM1 homo sapien |
| 13 | 134.5 | 9.7 | 704 13 Q90865 | Q90865 gallus galli |
| 14 | 134 | 9.6 | 716 11 Q91XG8 | Q91XG8 mus musculu |
| 15 | 131.5 | 9.4 | 313 13 Q9PU78 | Q9PU78 ctocodylus |
| 16 | 131 | 9.4 | 154 4 Q96SE8 | Q96SE8 homo sapien |

| | | | | |
|----|-------|-----|----------------|---------------------|
| 17 | 131 | 9.4 | 608 13 Q9PTW7 | Q9PTW7 struthio ca |
| 18 | 130.5 | 9.4 | 717 13 P70006 | P70006 xenopus lae |
| 19 | 129 | 9.3 | 616 6 Q97507 | Q97507 sus scrofa |
| 20 | 128.5 | 9.2 | 806 6 Q18783 | Q18783 macropus eu |
| 21 | 127 | 9.1 | 157 6 Q9TVA8 | Q9TVA8 bos taurus |
| 22 | 126 | 9.0 | 716 13 Q91691 | Q91691 xenopus lae |
| 23 | 125.5 | 9.0 | 560 4 Q14520 | Q14520 homo sapien |
| 24 | 125 | 8.8 | 728 6 Q9BR09 | Q9BR09 felis silve |
| 25 | 121.5 | 8.7 | 399 4 Q96GL8 | Q96GL8 homo sapien |
| 26 | 121.5 | 8.7 | 420 4 Q9BTE9 | Q9BTE9 papio hamed |
| 27 | 119 | 8.5 | 334 6 Q46507 | Q46507 papio hamed |
| 28 | 119 | 8.5 | 710 13 Q91402 | Q91402 xenopus lae |
| 29 | 117.5 | 8.4 | 812 11 Q91WJ5 | Q91WJ5 mus musculu |
| 30 | 114.5 | 8.2 | 113 4 Q9UIR7 | Q9UIR7 homo sapien |
| 31 | 114 | 8.2 | 113 4 Q9UIR5 | Q9UIR5 homo sapien |
| 32 | 114 | 8.2 | 202 13 Q90675 | Q90675 gallus galli |
| 33 | 114 | 8.2 | 452 13 Q90Y90 | Q90Y90 xenopus lae |
| 34 | 114 | 8.2 | 607 13 Q91001 | Q91001 gallus galli |
| 35 | 112.5 | 8.1 | 359 6 Q8MMR1 | Q8MMR1 canis famli |
| 36 | 112 | 8.0 | 567 4 Q13208 | Q13208 caenorhabdi |
| 37 | 112 | 8.0 | 37 6 P91823 | P91823 caenorhabdi |
| 38 | 112 | 8.0 | 1145 5 Q9BRL8 | Q9BRL8 alysia cal |
| 39 | 111.5 | 8.0 | 597 11 Q35727 | Q35727 mus musculu |
| 40 | 111 | 8.0 | 726 13 Q24488 | Q24488 gnosophilis |
| 41 | 110.5 | 7.9 | 685 5 Q24488 | Q24488 streptomyce |
| 42 | 110.5 | 7.9 | 2358 16 Q9ULV8 | Q9ULV8 streptomyce |
| 43 | 108 | 7.8 | 215 13 Q42341 | Q42341 gallus galli |
| 44 | 106 | 7.6 | 378 13 Q90WPO | Q90WPO trichemys s |
| 45 | 105.5 | 7.6 | 709 13 Q90ZNE | Q90ZNE brachydanto |

ALIGNMENTS

RESULT 1

ID Q96FE7 PRELIMINARY; PRT; 263 AA.
AC Q96FE7
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Unknown (protein for MGCL17330).
OS Homo sapiens (Human).
OC Mammalia; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euteleostomi; Primates; Catarrhini; Hominoidea; Homo.
RN NCBI_TaxID=9606;
RX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011049; AAH11049.1;
DR InterPro; IPR000001; Kinigle.
DR Pfam; PF00051; Kinigle; 1.
DR ProDom; PD000395; Kinigle; 1.
DR PROSITE; PS00021; KINIGLE_1; UNKNOWN_1.
DR PROSITE; PS50070; KINIGLE_2; 1.
SQ SEQUENCE 263 AA; 28234 MW; 197C3EEB888FA242 CRC64;

Query Match 100.0%; Score 1393; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 8.9e-118;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAWVOAFVSNMLAEAYSGGCGFMDNGHLVREDQTSAPAGRLCLMWLDAOSGLASAP 60
1 MLAWVOAFVSNMLAEAYSGGCGFMDNGHLVREDQTSAPAGRLCLMWLDAOSGLASAP 60
QY 61 VSGAGNHSYCRNDEDEPRGPGVCYSGAGVPEKRPCEBDRCEPTTSQALPFTTEIOEAS 120
61 VSGAGNHSYCRNDEDEPRGPGVCYSGAGVPEKRPCEBDRCEPTTSQALPFTTEIOEAS 120
QY 121 EGPAGDEVQVPAPANALPKRSSTAAVQPVIGISQRTPMNSKEXKULGTLGVVIGITMMVI 160
EGPAGDEVQVPAPANALPKRSSTAAVQPVIGISQRTPMNSKEXKULGTLGVVIGITMMVI 160

```

Db 121 EGPADENVQFAPANALPARSEAAAVDPVIGISQRVNMSSKEKDLGTLGVIGITMVI 180
Qy 181 IIAAGAILIGSYKRGKDLKQHDQKRCEREMQRTTLPLSAFTNPTCEIYDEKTVVHT 240
Db 181 IIAAGAILIGSYKRGKDLKQHDQKRCEREMQRTTLPLSAFTNPTCEIYDEKTVVHT 240
Qy 241 SQTPVDPQEGSTPLMGOAGTPGA 263
Db 241 SQTPVDPQEGSTPLMGOAGTPGA 263

RESULT 2
000318 PRELIMINARY; PRT; 263 AA.
ID 000318;
AC 000318;
DT 01-JUN-1997 (TREMblrel. 04, Created)
DT 01-JUN-1997 (TREMblrel. 04, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE MUSC.DJ51SN1.2 protein.
CN MUSC.DJ51SN1.2.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Du Z., Scheet P., Harper M.;
RT "The sequence of H. sapiens PAC clone RP3-51SN1."
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Materon R.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC002073; AAB54054.1; -.
DR HSSP: P00749; 1XDU.
DR InterPro: IPR000001; Kringle.
DR Pfam: PF00051; Kringle_1.
DR PRINTS: PR00018; KRINGLE; FALSE_NEG.
DR PRODOM: PD000395; Kringle; 1.
DR SMART: SM00130; KR_1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
SQ SEQUENCE 263 AA; 28248 MW; 197C3EBE854A242 CRC64;

Query Match 99.8%; Score 1390; DB 4; Length 263;
Best Local Similarity 99.6%; Pred. No. 1.7e-117;
Matches 262; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ENAMEL_ORGAN;
RA Ding Y., Xue J., Bartlett J.D.;
RT "T-plasminogen activator in tooth tissues."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF364605; AAM00297.1; -.
SQ SEQUENCE 562 AA; 63668 MW; F956B4C77CB101E8 CRC64;

Query Match 12.0%; Score 167; DB 6; Length 562;
Best Local Similarity 36.6%; Pred. No. 1.2e-06;
Matches 41; Conservative 12; Mismatches 33; Indels 26; Gaps 6;

Qy 25 CFMDNGHLYREDQTSPPAGLRCLNWLDAQSGIASAPVS-----GAGNHSYCRNPDE 75
Db 127 CYEDQGITRYKGTWSTESGAECVNM--NTSGIASPYNGRPRPAVYLGLGNHNYCRNPDK 184
Qy 76 DPGPWCYV-SGEAGVPEKRPCEDLRCPEPTTSQALPAFTTEIOASEGPGAD 126
Db 185 DSR-PWCYIFKAEKXSPD-----PC-----STPACTKEKECYTGKGLD 222

RESULT 4
09B099 PRELIMINARY; PRT; 516 AA.
ID 09B099;
AC 09B099;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Similar to plasminogen activator, tissue.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: BC002795; AAH02795.1; -.
DR HSSP: P00750; 1ASH.
DR MEROPS: S01.242; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00051; Kringle; 2.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRODOM: PD00018; KRINGLE.
DR PRODOM: PD000395; Kringle; 2.
DR SMART: SM00181; EGF_1.
DR SMART: SM00001; EGF-like; 1.
DR SMART: SM00130; KR_2.
DR SMART: SM00020; TRYPSIN; 1.
DR PROSITE: PS00022; EGF_2; UNKNOWN_1.
DR PROSITE: PS01166; EGF_2; UNKNOWN_2.
DR PROSITE: PS00021; KRINGLE_1; UNKNOWN_2.
DR PROSITE: PS00070; KRINGLE_2; 2.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.

```



```

DE Plasminogen precursor.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=LIVER;
RA Browne M.J., Chapman C.G., Dodd I., Carey J.E., Lawrence G.M.P.,
RA Mitchell D., Robinson J.H.;
RT "Expression of recombinant human plasminogen and aglycoplasminogen in
RT HeLa cells." 0:0-0(1993).
RT Fibrinolysis.
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE
CC EMIL; M74220; AAA6451.1; -.
DR HSSP; P00747; 2PK4.
DR MEROPS; S01.233; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 5.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; TRY_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS50070; KRINGLE_2; 5.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Serine protease; Signal.
KW SIGNAL.
FT CHAIN 1 19 POTENTIAL.
FT 20 810 PLASMINOGEN.
SQ SEQUENCE 810 AA; 90555 MW; B05C7DB0020B3C CRC64;

Query Match 10.2%; Score 142; DB 4; Length 810;
Best Local Similarity 37.24; Pred. No. 0.00035;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 25 CFWDNGHLYREDQTSFAPGLRCLNWLDA----QSGLASAPVSGAGNHSYCNPNDEDPG 79
DB 103 CKTGNGKNYRGITWSTKNGITCQKWSSTSPHRERFSPATHPEGL-ENNYCRNPDDPG 161
QY 80 PWCYVSGEAGVPEKR--PCEDLRCP 103
DB 162 PWCYTTD---PEKRYDYCDILECEE 183

RESULT 8
Q25101 PRELIMINARY; PRT; 385 AA.
AC Q25101;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Serine protease.
GN HMERP1
OS Herdmania momus.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Pyrosidae; Herdmania.
OX NCBI_TaxID=7735;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=LIVER;
RA Arnold U.W., Kennett C., Javin M.F.;
RA "Transgenic expression of a novel serine protease in the ectoderm of

```

```

RT the ascidian Herdmania momus during development.";
RL Dev. Genes Evol. 206:455-463(1997).
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC EMBL; U63117; AAB6650.1; -.
DR HSSP; P00763; IDPO.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR Pfam; PF00024; PAN; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRY_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 385 AA; 42935 MW; BFB1D05D5232E6A0 CRC64;

Query Match 10.0%; Score 139.5; DB 5; Length 385;
Best Local Similarity 20.2%; Pred. No. 0.00023;
Matches 70; Conservative 47; Mismatches 127; Indels 103; Gaps 14;

QY 2 LLA--VOAFVSNMILABAYSGGCF-WDNGHLYREDQTSFAPGLRCLNWLDAQS---- 54
DB 7 LVIWIIINGFVESH-----SECFDIENPESYOGALSRLGGETQSW-DLQTPHNG 56
QY 55 --GLASAPVSGAGNHSYCNPNDEDPGWCYVSGE-----AGVPEKRPEDEDR 100
DB 57 KITSNENPENGAGNYYCNPNQDNGRCWCYTNENFMRYCDIPICSNPPVTLPSIE 116
QY 101 CETH--SQALPAFTTEQESSECP---GADENQ----- 129
DB 117 CKRTPELPDPTGTYGLDKSRAKTNELHIVGTTVTHGSIPIQVSLAKELHFGGSI 176
QY 130 ----VFAPNALPANSAAVQPIYGISQVRNRSKEKKDGLTGYVL----- 173
DB 177 LNRNWLTPAHCTRKQOPKQKTLALGDYDRIQDSEKK-----VGRLLFNHEKTPAT 232
QY 174 --GITMVIITAIIGAGILGYSYRKGDKQHDQVC-----EREMQRITL 218
DB 233 FENDITLTKMDTISITATIFQGSVFPANRVPAAKSIIVSGMDPTKGTQDVKLQVTL 292
QY 219 PLSAFTNPCEIYDEKTV---VHTSOTPVDPQEGSTPLMGQACTP 261
DB 293 PWSFP--KLCKKLKYVGAVALVFETSLCAAYKKGKDSGQDSGSP 317

RESULT 9
P70521 PRELIMINARY; PRT; 716 AA.
AC P70521;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Macrophage stimulating protein precursor.
GN MSP
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=LIVER;
RC MEDLINE=97011126; PubMed=8858136;
RA Ohsuhiro K., Iwama A., Matsuno K., Ezaki T., Sakamoto O., Hamaguchi I.,
RA Takasu N., Suda T.;
RT "Molecular cloning of Rat Macrophage-stimulating protein and its

```

RT involvement in the Male Reproductive System.";
 RL Biochem. Biophys. Res. Commun. 227:273-280(1996).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL; X95096; CA64473.1; -.
 DR HSSP; P00747; 1KRN.
 DR MEROPS; S01.975; -.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringles.
 DR InterPro; IPR003609; Pan. app.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00051; Kringles; 4.
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00083; Trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00181; KRINGLE.
 DR ProDom; PD000395; Kringles; 4.
 DR SMART; SM00130; KR; 4.
 DR SMART; SM00473; PAN_AP; 1.
 DR SMART; SM00020; TRYPSIN; 1.
 DR PROSITE; PS00021; KRINGLE_1; 4.
 DR PROSITE; PS50070; KRINGLE_2; 4.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR Hydrolase; Serine protease; Signal.
 KW SIGNAL.
 FT SIGNAL.
 SQ SEQUENCE 716 AA; 80733 MW; 06B7DF3EF56D921F CRC64;
 Query Match 9.8%; Score 137; DB 11; Length 716;
 Best Local Similarity 27.3%; Pred. No. 0.00084;
 Matches 48; Conservative 12; Mismatches 56; Indels 60; Gaps 8;
 Y 2 ILAVQALVLSGNMLAEVSGS-----GCFMDNGHYREDQTSAPGLRCLAM 49
 DB 80 LRPWQ-----HSLRAQLHSSLCPLPQKDYATCIMNGASVGTARFADLPCQAM 134
 Y 50 ---LDAOSGLASAPVSGAGSHVSCNPNDEPRGMCVVS----- 65
 DB 135 SRFPNDKHYTPYKNGL-ENFCNPNFGDPRGMCYTTNRSVRQSGCISCREANCVW 193
 Y 86 -----GEAGVPER-RPCE--DLACPET-----TSQALPAFTTEIOASGSP 123
 DB 194 CNGEDYRGEDVLTSGRECORDLPHSHFPEKFPDALKNDYCNPDASBP 249
 RESULT 10
 Q90504 PRELIMINARY; PRT; 420 AA.
 ID Q90504;
 AC Q90504;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Thrombin.
 OS Eptatretus scutell (Pacific hagfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniformes;
 OC Myxiniidae; Eptatrelinae; Eptatretus.
 OC NCBI_TaxID=7765;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RC MEDLINE=92212913; PubMed=1557383;
 RA Banfield D.K., Macgillivray R.T.;
 RT "Partial characterization of vertebrate prothrombin cDNAs:
 RT amplification and sequence analysis of the B chain of thrombin from
 RT rife different species.";
 RI Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
 RP [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RC MEDLINE=94223694; PubMed=7513365;
 RA Banfield D.K., Irwin D.M., Walz D.A., Macgillivray R.T.;
 RT "Evolution of prothrombin: isolation and characterization of the cDNAs
 RT encoding chicken and hagfish prothrombin.";

RL J. Mol. Evol. 38:177-187(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Banfield D.K.;
 RL Submitted (DEC-1991) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL; M81393; AAA21620.1; -.
 DR HSSP; P00734; 1UVS.
 DR MEROPS; S01.217; -.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringles.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00051; Kringles; 1.
 DR Pfam; PF00083; Trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00181; KRINGLE.
 DR ProDom; PD000395; Kringles; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRYPSIN; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Hydrolase; Serine protease.
 KW SEQUENCE 420 AA; 47888 MW; 64522AA21A57B67A CRC64;
 Query Match 9.8%; Score 136.5; DB 13; Length 420;
 Best Local Similarity 26.5%; Pred. No. 0.00047;
 Matches 39; Conservative 16; Mismatches 77; Indels 15; Gaps 4;
 Y 25 CPMNDNHLVREDQTSAPGLRCLMWDASGLASAPVSGAG-NHSYCRPNDEPRGMCY 83
 DB 17 CVRREGSDYRGDLMTWTGKPLCPNMGSYNLPDSFTTAGLTSYCNRPDSDSGWCY 76
 Y 84 VSGEAVP-----EKRPEDLRCPETTSQALPAFTTEIOASGSGADEVOVAPANLPA 139
 DB 77 TKVEGTVDVYQALNCE-----SGDFFVGTIDVQLSGRSGAAETLFFPNKTFGN 129
 Y 140 RSPAAVDPVIGIGISQVRNMSKEKDL 166
 DB 130 GEEBCGRMPBLQK---NDRSEDL 153
 RESULT 11
 Q90503 PRELIMINARY; PRT; 812 AA.
 ID Q90503;
 AC Q90503;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Plasminogen protein precursor (EC 3.4.21.7).
 GN PLASMINOGEN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RC Banget K., Johnsen A.H., Thorsen S.;
 RT "Rat plasminogen cDNA and gene structure.";
 RT Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
 RP [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RC MEDLINE=9150378; PubMed=1645711;
 RA Kanalas U.O., Maket S.R.;
 RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
 RT receptor site for plasminogen.";
 RL J. Biol. Chem. 266:10825-10829(1991).

```

CC -1- SIMILARITY. BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: A1242649; CAB46014.1; -.
DR HSSP: P00747.1PKK.
DR MEROPS: S01.233.
DR InterPro: IPR01314; Chymotrypsin.
DR InterPro: IPR00001; Kringle.
DR InterPro: IPR00304; PAN.
DR InterPro: IPR00369; Pan_app.
DR InterPro: IPR01254; Ser_protease_Try.
DR InterPro: IPR01400; Somatotropin.
DR Pfam: PF00051; Kringle; 5.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 5.
DR SMART: SM00130; KR; 4.
DR SMART: SM00473; PAN_Ap; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS00021; KRINGLE_1; 5.
DR PROSITE: PS00070; KRINGLE_2; 5.
DR PROSITE: PS00338; SOMATOTROPIN_2; UNKNOWN_1.
DR PROSITE: PS02040; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
DR Hydroxylase; Serine protease; Signal.
FT SIGNAL 1
FT CHAIN 1
SQ SEQUENCE 812 AA; 90535 MW; 8C703C51410BEC9E CRC64;

Query Match 9.7%; Score 135; DB 11; Length 812;
Best Local Similarity 30.9%; Pred. No. 0.0015;
Matches 33; Conservative 14; Mismatches 48; Indels 14; Gaps 5;

QY 25 CFPMNGHLREDQTSFAPGLCLNWLDA-----QSLASAVSGAGNHSTCRNPEDPRG 79
DB 376 CYQNGSGYRGTSSITNTGKCSQSMVMTPHSHSKTPANFPDACL-EMNYCRNPDDRG 434
QY 80 PWCYVSGEACVPEKR--PCEDLRCPETTSQALPAFTTEIQASGPGADE 127
DB 435 PWCFTTD---PSVMEYCNLKRCEGTGGV--AESAIIVQVPSAPGTS 478

RESULT 12
Q98ZM1 PRELIMINARY; PRT; 395 AA.
AC Q98ZM1;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Neocatal chymolytic agent alpha-form (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_Taxid=9606;
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dou D.;
RT "A brain-type plasminogen activator."
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY. BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: AF060825; AAK1956.1; -.
DR HSSP: P00750.1PK2.
DR MEROPS: S01.233.
DR InterPro: IPR01314; Chymotrypsin.
DR InterPro: IPR00003; Fibrinol.
DR InterPro: IPR00001; Kringle.
DR InterPro: IPR01254; Ser_protease_Try.
DR Pfam: PF00039; Enl; 1.
DR Pfam: PF00051; Kringle; 1.
DR Pfam: PF00089; trypsin; 1.

```

```

DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 1.
DR SMART: SM00056; ENL; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS01253; FIBROECTIN_1.1.
DR PROSITE: PS00021; KRINGLE_1; UNKNOWN_1.
DR PROSITE: PS00070; KRINGLE_2; 1.
DR PROSITE: PS00240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
DR Hydroxylase; Serine protease.
FT NON_TER 395
SQ SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match 9.7%; Score 134.5; DB 4; Length 395;
Best Local Similarity 32.3%; Pred. No. 0.00066;
Matches 32; Conservative 13; Mismatches 33; Indels 21; Gaps 4;

QY 4 AMVAFIVSNML-----LAAYSGGCFWVNGHLYREDQTSFAPGLCLNWL-----49
DB 18 SWLPVLRSRVRYECWNSGRACSEGNSDCYFGNSAYRGTHSLTSGASCLPMNSML 77
QY 50 ----LDQSGLASAPVSGAGNHSTCRNPEDPRGPGCVY 84
DB 78 IGKVTNQN--PEAQALGIGHNYCRNPDDAK-PWCHV 113

RESULT 13
Q90865 PRELIMINARY; PRT; 704 AA.
AC Q90865;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Hepatocyte growth factor-like/macrophage stimulating protein.
GN HGF/MSP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_Taxid=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSEU=LIVER;
RC MEDLINE=96029010; PubMed=7554499;
RA Thery C., Sharpe M.J., Bailey S.J., Stern C.D., Gherardi E.;
RT "Expression of HGF/SF, HGF1/MSP and c-met suggests new functions
RL during early chick development."
RL Dev. Genet. 17:90-101(1995).
CC -1- SIMILARITY. BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: X84043; CAAS8662.1; -.
DR HSSP: P00747; ICEA.
DR MEROPS: S01.977; -.
DR InterPro: IPR01314; Chymotrypsin.
DR InterPro: IPR00001; Kringle.
DR InterPro: IPR00304; PAN.
DR InterPro: IPR00369; Pan_app.
DR InterPro: IPR01254; Ser_protease_Try.
DR Pfam: PF00051; Kringle; 4.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 4.
DR SMART: SM00130; KR; 4.
DR SMART: SM00473; PAN_Ap; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS00021; KRINGLE_1; 4.
DR PROSITE: PS00070; KRINGLE_2; 4.
DR PROSITE: PS02040; TRYPsin_DOM; 1.

```


KW Hydrolase; Serine protease.
 SQ SEQUENCE 704 AA; 79341 MW; CAB0DBCC41367C37 CRC64;
 Query Match 9.7%; Score 134.5; DB 13; Length 704;
 Best Local Similarity 28.8%; Pred. No. 0.0014;
 Matches 32; Conservative 11; Mismatches 45; Indels 23; Gaps 4;

QY 25 CFWDNGHLVREDQTSPPAPGLRCLNWLDAQSGLASAP-----VSGAGNHSYCRNPDED 76
 DB 108 CIYANGETSRGRDITRERGRCHM-----QATTPHDIRFLPSLANGLEENYCRNPDRD 161

QY 77 PRGPMCV-----SGAGVPEKPCEDLRCPETTSQALPAFTTEIQEASE 121
 DB 162 KRGPWCYTVDPNVRHOSCGI---KKCEDAVCMTCNGEDYRGFVDHTSGTE 209

RESULT 14
 Q91XG8 PRELIMINARY; PRT; 716 AA.
 AC Q91XG8;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Hepatocyte growth factor-like.
 GN HGFL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC010551; AAH10551.1; -.
 DR MGD; MG1:96080; HGFL.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF00051; kringle; 4.
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR ProDom; PD000395; Kringle; 4.
 DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_4.
 DR PROSITE; PS50070; KRINGLE_2; 4.
 DR PROSITE; PS50240; TRYPsin_DOM; 1.
 DR Hydrolase; Serine protease.
 KW SEQUENCE 716 AA; 80693 MW; 12474C48A7D4B46D CRC64;
 SQ

Query Match 9.6%; Score 134; DB 11; Length 716;
 Best Local Similarity 28.2%; Pred. No. 0.0016;
 Matches 42; Conservative 9; Mismatches 46; Indels 52; Gaps 7;

QY 2 LLAWVQAFLVSNMLLAAYSGS-----GCFWDNGHLVREDQTSPPAPGLRCLNW 49
 DB 80 LLPWTQ-----HSLHTQLYHSSLCHLFOKKDYVRTCIMNGVSYRGVARTAGLPCQAW 134

QY 50 ---LDAQSGLASAPVSGAGNHSYCRNPDEDPRGPMCVS----- 85
 DB 135 SRFPNDHKYTPTRKNGI--EENFCRNPDPGDPGRGWCYTTNRSVRFOSCGIKTCREAVCVL 193

QY 86 -----GEAGVPEK-PPCE--DURCPET 104
 DB 194 CNGEDYRGFVDHTSGRECRQMDLQHPHS 222

RESULT 15
 Q9PU78 PRELIMINARY; PRT; 313 AA.
 AC Q9PU78;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE Hepatocyte growth factor-like protein (Fragment).
 OS Crocodylus niloticus (Nile crocodile) (African crocodile).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Crocodylidae; Crocodylinae; Crocodylus.
 OX NCBI_TaxID=8501;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=20022983; PubMed=10555283;
 RA Hughes S., Zelus D., Mouchiroud D.;
 RT "Warm-blooded isochore structure in Nile crocodile and turtle.";
 RL Mol. Biol. Evol. 16:1521-1527(1999).
 DR EMBL; AJ011396; CAB56422.1; -.
 DR HSSP; P00747; IHPJ.
 DR MEROPS; S01.977; -.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF00051; kringle; 2.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPsin_DOM; 1.
 KW Hydrolase; Serine protease.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 313 AA; 34793 MW; 8E084704958B5AA2 CRC64;

Query Match 9.4%; Score 131.5; DB 13; Length 313;
 Best Local Similarity 29.8%; Pred. No. 0.0092;
 Matches 28; Conservative 11; Mismatches 42; Indels 13; Gaps 4;

QY 25 CFWDNGHLVREDQTSPPAPGLRCLNWLDAQSGLAS-AYVSGAGN-----SYCRNPDEDPRG 80
 DB 30 CYHNGELVYRGHTSKTRKGVTCQKMSQSPRVQISPTTHPAHLADENYCRNPNDSHGP 89

QY 81 WCYVSGE-----AGVPEKPCEDLRCPETTSQA 108
 DB 90 WCYTMDPRTPPDYCGI---KPCAGDKLPSVLENA 120

Search completed: April 7, 2003, 07:06:00
 Job time : 59.4987 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 06:41:55 ; Search time 41.3589 Seconds
(without alignments)
847.337 Million cell updates/sec

Title: US-10-057-951-2
 Perfect score: 1393
 Sequence: 1 MLAWVQAFIVSNMLAEAY.....PVDPQEGSTPLMGAGTPGA 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Genseq.101002.*

1: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1980.DAT.*

2: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1981.DAT.*

3: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1982.DAT.*

4: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1983.DAT.*

5: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1984.DAT.*

6: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1985.DAT.*

7: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1986.DAT.*

8: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1987.DAT.*

9: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1988.DAT.*

10: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1989.DAT.*

11: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1990.DAT.*

12: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1991.DAT.*

13: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1992.DAT.*

14: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1993.DAT.*

15: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1994.DAT.*

16: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1995.DAT.*

17: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1996.DAT.*

18: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1997.DAT.*

19: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1998.DAT.*

20: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1999.DAT.*

21: /SIDS2/gcgdata/genseq/genseqp-emb1/AA2000.DAT.*

22: /SIDS2/gcgdata/genseq/genseqp-emb1/AA2001.DAT.*

23: /SIDS2/gcgdata/genseq/genseqp-emb1/AA2002.DAT.*

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|-----------|----------------------|
| 1 | 1393 | 100.0 | 263 | 20 | AAY05219 | Kringel protein s |
| 2 | 1389 | 100.0 | 263 | 20 | AAW87769 | Human tissue plasmin |
| 3 | 1383 | 100.0 | 263 | 22 | AAR00300 | Human tissue plasmin |
| 4 | 1350 | 99.8 | 263 | 23 | AFB06149 | Human CR2-like polyA |
| 5 | 1364 | 99.4 | 263 | 21 | ABZ43237 | Human ORX-ORF30U |
| 6 | 1379 | 99.0 | 263 | 22 | AAW93748 | Human polypeptide |
| 7 | 1167.5 | 83.8 | 286 | 20 | AAV05220 | Kringel protein s |
| 8 | 322 | 23.1 | 66 | 22 | ABB37905 | peptide #541l enco |
| 9 | 322 | 23.1 | 66 | 22 | AAW823159 | protein #518 enco |
| 10 | 322 | 23.1 | 66 | 22 | AAW58537 | Human brain expres |

| | | | | | | |
|----|-------|------|-----|----|----------|---------------------|
| 1 | 322 | 23.1 | 66 | 22 | AAW71037 | Human bone marrow |
| 2 | 322 | 23.1 | 66 | 22 | AAW18900 | Peptide #5234 enco |
| 3 | 322 | 23.1 | 66 | 22 | AAW31314 | Peptide #5351 enco |
| 4 | 322 | 23.1 | 66 | 23 | ABC40828 | Human peptide enco |
| 5 | 305 | 22.0 | 56 | 20 | AAW12615 | Human 5' EST seque |
| 6 | 305 | 22.0 | 55 | 20 | AAW12397 | Human 5' EST seque |
| 7 | 196 | 14.1 | 39 | 19 | AAW72641 | Nervous glia cell |
| 8 | 192 | 13.8 | 39 | 19 | AAW72640 | Nervous glia cell |
| 9 | 160.5 | 11.5 | 527 | 13 | AAW20217 | t-PA analogue expr |
| 10 | 160.5 | 11.5 | 527 | 13 | AAW20220 | t-PA analogue expr |
| 11 | 160.5 | 11.5 | 527 | 13 | AAW20218 | t-PA analogue expr |
| 12 | 160.5 | 11.5 | 527 | 13 | AAW20219 | t-PA analogue expr |
| 13 | 160.5 | 11.5 | 593 | 8 | AAW70059 | Sequence of hybrid |
| 14 | 160.5 | 11.5 | 650 | 8 | AAW70060 | Sequence of hybrid |
| 15 | 160 | 11.5 | 351 | 20 | AAW25407 | Human tissue coding |
| 16 | 159.5 | 11.5 | 472 | 10 | AAW94416 | Sequence of coding |
| 17 | 159.5 | 11.5 | 527 | 13 | AAW20223 | t-PA analogue expr |
| 18 | 159.5 | 11.5 | 527 | 13 | AAW20222 | t-PA analogue expr |
| 19 | 159.5 | 11.5 | 527 | 13 | AAW20221 | t-PA analogue expr |
| 20 | 157.5 | 11.3 | 527 | 13 | AAW54154 | t-PA mutant (N142S |
| 21 | 157.5 | 11.3 | 527 | 13 | AAW54157 | t-PA mutant (N142S |
| 22 | 157.5 | 11.3 | 562 | 10 | AAW94379 | Amino acid sequenc |
| 23 | 157.5 | 11.3 | 562 | 10 | AAW94380 | Amino acid sequenc |
| 24 | 156.5 | 11.2 | 439 | 16 | AAW68851 | Delta 2-69 tissue |
| 25 | 156.5 | 11.2 | 439 | 16 | AAW70889 | Human tissue PA va |
| 26 | 156.5 | 11.2 | 483 | 16 | AAW70877 | Human tissue PA va |
| 27 | 156.5 | 11.2 | 483 | 16 | AAW70878 | Human tissue PA va |
| 28 | 156.5 | 11.2 | 483 | 16 | AAW70879 | Human tissue PA va |
| 29 | 156.5 | 11.2 | 483 | 16 | AAW70880 | Human tissue PA va |
| 30 | 156.5 | 11.2 | 483 | 16 | AAW70881 | Human tissue PA va |
| 31 | 156.5 | 11.2 | 483 | 16 | AAW70882 | Human tissue PA va |
| 32 | 156.5 | 11.2 | 483 | 16 | AAW70883 | Human tissue PA va |
| 33 | 156.5 | 11.2 | 483 | 16 | AAW70884 | Human tissue PA va |
| 34 | 156.5 | 11.2 | 483 | 16 | AAW70885 | Human tissue PA va |
| 35 | 156.5 | 11.2 | 483 | 16 | AAW70886 | Human tissue PA va |

ALIGNMENTS

| | |
|----------|---|
| RESULT 1 | |
| AAV05219 | |
| ID | AAV05219 standard; Protein; 263 AA. |
| XX | |
| AC | AAV05219; |
| XX | |
| DT | 17-JUN-1999 (first entry) |
| XX | |
| DE | Kringle1 protein sequence. |
| XX | |
| KW | Kringle1; cancer; inflammation; autoimmunity; rheumatoid arthritis; AIDS; |
| KW | CNS inflammation; cerebellar degeneration; Alzheimer's disease; asthma; |
| KW | Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis; |
| KW | neurological abnormality; ischaemia reperfusion injury; ischaemic injury; |
| KW | cardiovascular disease; kidney disease; liver disease; aplastic anaemia; |
| KW | myocardial infarction; hypotension; hypertension; allergy; infection; |
| KW | myelodysplastic syndrome; haematologic abnormality; diagnosis; therapy; |
| KX | male pattern baldness. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | W09911788-A1. |
| XX | |
| PD | 11-MAR-1999. |
| XX | |
| PF | 02-SEP-1998; 98WO-US18270. |
| XX | |
| PR | 01-SEP-1998; 98US-0144889. |
| PR | 02-SEP-1997; 97US-0056032. |
| XX | |
| PA | (SMIK) SMITHKLINE BEECHAM CORP. |
| XX | |
| PI | Albone EF, Kikly KK; |

XX WPI: 1999-214707/18.
 DR N-PSDB; AAX28354.
 XX
 FT New kringle1 polypeptides and polynucleotides
 XX
 PS Claim 1; Page 31-32; 42pp; English.
 XX
 CC This sequence is a Kringle1 polypeptide of the invention.
 CC The kringle1 polypeptides (I) are used to screen for agonists and
 CC antagonists. Agonists are used to treat subjects in need of enhanced
 CC activity or expression of (I). Antagonists are used to treat subjects
 CC having need to inhibit the activity or expression of (I). The methods can
 CC be used to treat conditions such as cancer, inflammation, autoimmunity,
 CC allergy, asthma, rheumatoid arthritis, CNS inflammation, cerebellar
 CC degeneration, Alzheimer's disease, Parkinson's disease, multiple
 CC sclerosis, amyotrophic lateral sclerosis, head injury damage and other
 CC neurological abnormalities, ischemia reperfusion injury, cardiovascular
 CC disease, kidney disease, liver disease, ischemic injury, myocardial
 CC infarction, hypotension, hypertension, AIDS, myelodysplastic syndromes
 CC and other haematologic abnormalities, aplastic anaemia, male pattern
 CC baldness, and bacterial, fungal, protozoan and viral infections. The
 CC kringle1 polypeptides may also be used to generate antibodies.
 CC Determining the presence or absence of mutations in, and analysing for
 CC the presence or absence of expression of, kringle1 polynucleotides can be
 CC used to diagnose a disease or susceptibility to a disease related to
 CC expression or activity of kringle1 proteins. The polynucleotides may also
 CC be used for chromosome identification, and mapping.
 XX
 SQ Sequence 263 AA;
 Query Match 100.0%; Score 1393; DB 20; Length 263;
 Best Local Similarity 100.0%; Pred No. 9e-120;
 Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLAMVQARVSNMLAEVSGGCFMNGHTRFEDTSPAPRLCLMWLDNOSGLASAP 60
 DB 1 MLAMVQARVSNMLAEVSGGCFMNGHTRFEDTSPAPRLCLMWLDNOSGLASAP 60
 QY 61 VSGAGNHSYCRNPDEDPKPGWCYSGEAGVEKAPCEDLRCPETTSQALPAFTTEIOEAS 120
 DB 61 VSGAGNHSYCRNPDEDPKPGWCYSGEAGVEKAPCEDLRCPETTSQALPAFTTEIOEAS 120
 QY 121 EGPAGDEVQVFAPANALPARSEAAVQPVIGISQRYVMSKEKQDGLGYVLGITMVI 180
 DB 121 EGPAGDEVQVFAPANALPARSEAAVQPVIGISQRYVMSKEKQDGLGYVLGITMVI 180
 QY 181 IIAIGAGIILGYSKRGKQDKCEHMRITLPLSAFTNPCEIYDEKTVVYHT 240
 DB 181 IIAIGAGIILGYSKRGKQDKCEHMRITLPLSAFTNPCEIYDEKTVVYHT 240
 QY 241 SQRTVPDQEGSTPLMGQACTPCA 263
 DB 241 SQRTVPDQEGSTPLMGQACTPCA 263
 RESULT 2
 AAM87769
 ID AAM87769 standard; Protein; 263 AA.
 AC AAM87769;
 XX 29-MAR-1999 (first entry)
 DE Human tissue plasminogen activator-like protease t-PALP.
 KW Tissue plasminogen activator-like protease; t-PALP; human;
 KW circulatory system-related disorder; blood clotting; stroke;
 KW thrombosis; peripheral arterial occlusion; pulmonary embolism;
 KW myocardiothrombosis; diagnosis; therapy.
 XX
 OS Homo sapiens.
 XX

PH Key Location/Qualifiers
 FT Peptide 1..21
 FT Peptide /label= Sig_peptide
 FT Protein 22..263
 FT Domain /label= Mac protein
 FT Domain 25..84
 FT Domain /note= "kringle domain"
 FT Domain 85..263
 FT Peptide /note= "protease domain"
 FT Peptide 22..31
 FT Peptide /note= "epitope-bearing region"
 FT Peptide 35..44
 FT Peptide /note= "epitope-bearing region"
 FT Peptide 71..81
 FT Peptide /note= "epitope-bearing region"
 FT Peptide 91..107
 FT Peptide /note= "epitope-bearing region"
 FT Peptide 119..128
 FT Peptide /note= "epitope-bearing region"
 FT Peptide 138..147
 FT Peptide /note= "epitope-bearing region"
 FT Peptide 155..167
 FT Peptide /note= "epitope-bearing region"
 FT Peptide 193..203
 FT Peptide /note= "epitope-bearing region"
 FT Peptide 206..215
 FT Peptide /note= "epitope-bearing region"
 FT Peptide 227..237
 FT Peptide /note= "epitope-bearing region"
 FT Peptide 243..252
 FT Peptide /note= "epitope-bearing region"
 PN W09854199-A1.
 PD 03-DEC-1998.
 PP 27-MAY-1998; 98MO-US10728.
 PR 28-MAY-1997; 97US-0048000.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 PI Ebner R, Moore PA, Ruben SM;
 DR WPI: 1999-070207/06.
 DR N-PSDB; AAV99636.
 XX
 PT New tissue plasminogen activator-like protease - useful in the
 PT diagnosis and treatment of circulatory system-related disorders
 XX
 XX Claim 1; Page 56-57; 76pp; English.
 XX
 CC This is the amino acid sequence of tissue plasminogen activator-like
 CC protease (t-PALP), a novel member of the serine protease family
 CC that shares sequence homology to human tissue plasminogen activator
 CC (see AAM87770). The t-PALP sequence was deduced from a cDNA clone
 CC (see AAV99636) derived from activated monocytes. The 2.5 kb t-PALP
 CC message has also been detected in heart, brain, lung, placenta,
 CC liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate,
 CC testis, ovary, small intestine, colon and peripheral blood
 CC leukocytes. Isolated nucleic acids encoding amino acids -21 to
 CC 242, -30 to 242, 1-242, 4-63 (kringle domain) and 64-242 (protease
 CC domain) of t-PALP or encoding epitope-bearing portions of t-PALP,
 CC are also claimed, as are recombinant vectors, host cells, and
 CC methods for producing t-PALP polypeptides. t-PALP may be used to
 CC detect and treat disorders related to the circulatory system, and
 CC to identify agonists and antagonists of t-PALP activity. The
 CC homology between t-PALP and tPA indicates that t-PALP may be
 CC involved in the regulation of normal and abnormal clotting
 CC in e.g. stroke, deep-vein thrombosis, peripheral arterial
 CC occlusion, pulmonary embolism and myocardiothrombosis.
 CC
 XX
 SQ Sequence 263 AA;

```

Query Match      100.0%; Score 1393; DB 20; Length 263;
Beat Local Similarity 100.0%; Pred. No. 9e-120;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAMVQAFIVSNMLLAEEVSGSGCFMNGHLFREQOTSPPAQGRCLNWLDAQSGLASAP 60
    |||||
DB 1 MLAMVQAFIVSNMLLAEEVSGSGCFMNGHLFREQOTSPPAQGRCLNWLDAQSGLASAP 60
    |||||

QY 61 VSGAGNSYCRNPDEDPBPCWCVSGEAGYEPKRPCEDEJRCETTSOLAPAFTEIOGAS 120
    |||||
DB 61 VSGAGNSYCRNPDEDPBPCWCVSGEAGYEPKRPCEDEJRCETTSOLAPAFTEIOGAS 120
    |||||

QY 121 BPGGADGVQVPAFANALPARESAANVQVIGISQVRNMSKKKDLGLIGVIGIMNT 180
    |||||
DB 121 BPGGADGVQVPAFANALPARESAANVQVIGISQVRNMSKKKDLGLIGVIGIMNT 180
    |||||

QY 181 IIAIGAGIILIGSYKRGKDKLEQHDQKVCEREMQRTLLPLAFNPTCEIYDEKTVVHT 240
    |||||
DB 181 IIAIGAGIILIGSYKRGKDKLEQHDQKVCEREMQRTLLPLAFNPTCEIYDEKTVVHT 240
    |||||

QY 241 SQTVPDQEGSTPLMGQAGTPGA 263
    |||||
DB 241 SQTVPDQEGSTPLMGQAGTPGA 263
    |||||

RESULT 3
AAE00300
ID AAE00300 standard; Protein; 263 AA.
XX
AC AAE00300;
XX
DT 13-JUN-2001 (first entry)
XX
DE Human tissue-plasminogen activator-like protease (t-PALP).
XX
KW Human; tissue-plasminogen activator-like protease; t-PALP;
KW therapy; vascular disease; stroke; deep vein thrombosis; keloid; asthma;
KW arterial occlusion; blood coagulation disorder; cerebroprotective;
KW autoimmune system disorder; human immunodeficiency syndrome; cyclostatic;
KW rheumatoid arthritis; graft-versus-host disease; chyloditis; candida;
KW insulin dependent diabetes; inflammatory eye disease; allergy; antiviral;
KW cardiovascular disease; heart disease; arhythmia; myocardial ischemia;
KW hyperproliferative disorder; hypertrophic scar; neurological disease;
KW Creutzfeldt-Jakob syndrome; neurodegenerative disorder; antiinflammatory;
KW Alzheimer's disease; Parkinson's disease; immunosuppressive; antifungal;
KW infectious disease; drug screening; gene therapy; neuroprotective;
KW cancer; ophthalmological; antibacterial; vulnerary.
XX
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX Binding-site 1..165
XX /note= "Binds to FLAG polypeptide to form
XX t-PALP-FLAG fusion protein"
XX Peptide 1..21
XX /label= Signal_peptide
XX Domain 4..63
XX /label= Kringle_domain
XX Region 12..21
XX /note= "Conserved region"
XX Protein 22..263
XX /note= "human mature tissue-plasminogen activator-like
XX t-PALP-FLAG fusion protein"
XX Region 22..38
XX /note= "Conserved region"
XX Region 22..31
XX /note= "Epitope-bearing portion"
XX Region 35..44
XX /note= "Epitope-bearing portion"
XX Region 39..49
XX /note= "Conserved region"

```

```

FT Region 50..63
FT /note= "Conserved region"
FT Region 63..84
FT /note= "Conserved region"
FT Domain 64..242
FT /label= Protease_domain
FT Region 71..81
FT /note= "Epitope-bearing portion"
FT Region 85..97
FT /note= "Conserved region"
FT Region 91..107
FT /note= "Epitope-bearing portion"
FT Region 100..118
FT /note= "Conserved region"
FT Region 119..128
FT /note= "Epitope-bearing portion"
FT Region 119..127
FT /note= "Conserved region"
FT Region 128..143
FT /note= "Conserved region"
FT Region 138..147
FT /note= "Epitope-bearing portion"
FT Region 146..163
FT /note= "Conserved region"
FT Region 155..167
FT /note= "Epitope-bearing portion"
FT Region 164..180
FT /note= "Conserved region"
FT Region 186..200
FT /note= "Conserved region"
FT Region 193..203
FT /note= "Epitope-bearing portion"
FT Region 201..220
FT /note= "Conserved region"
FT Region 206..215
FT /note= "Epitope-bearing portion"
FT Region 221..236
FT /note= "Conserved region"
FT Region 227..237
FT /note= "Epitope-bearing portion"
FT Region 237..248
FT /note= "Conserved region"
FT Region 243..252
FT /note= "Epitope-bearing portion"
FT Region 249..263
FT /note= "Conserved region"
XX
XX MO200125252-A1.
XX
XX 12-APR-2001.
XX
XX 03-OCT-2000; 2000WO-US27239.
XX
XX 04-OCT-1999; 99US-0411977.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Moore PA, Ruben SM, Ebner R;
XX
XX WPI; 2001-235402/24.
XX
XX N-PSDB; AAD03460.
XX
XX
XX New (gene encoding and antibody immunospecific for a)
XX tissue-plasminogen activator-like protease, useful for the diagnosis
XX and treatment of (cardio)vascular diseases, hyperproliferative
XX disorders, immune system disorders and cancers -
XX
XX Claim 17; Fig 1; 323pp; English.
XX
XX The present amino acid sequence is HWS1R42 clone human
XX tissue-plasminogen activator-like protease (t-PALP). The t-PALP
XX sequence and their (ant)agonists are useful for the diagnosis and
XX treatment of vascular diseases e.g. stroke, deep vein thrombosis and

```

CC arterial occlusion, blood coagulation disorders, (auto)immune system
 CC disorders e.g. human immunodeficiency syndrome, rheumatoid arthritis,
 CC graft-versus-host disease, thyroiditis, insulin dependent diabetes and
 CC inflammatory eye disease, allergic reactions e.g. asthma, cardiovascular
 CC diseases e.g. heart disease, arrhythmia and myocardial ischemia,
 CC hyperproliferative disorders, cancers, hypertrophic scars and keloids,
 CC neurological diseases e.g. Creutzfeldt-Jakob syndrome, neurodegenerative
 CC disorders e.g. Alzheimer's disease and Parkinson's disease and infectious
 CC disease e.g. viral, bacterial and fungal infections. The t-RAP sequences
 CC are also useful for drug screening. The t-RAP nucleotides are useful as
 CC chromosome markers and are involved in gene therapy.

XX Sequence 263 AA:

Query Match 100.0%; Score 1393; DB 22; Length 263;
 Best Local Similarity 100.0%; Pred. No. 9e-120;
 Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLAMVQAFIVSNMLIAEYSGGCFWMDNGHLTYREDQTSPPAGLRCLNMTDAQSGLASAP 60
 DB 1 MLAMVQAFIVSNMLIAEYSGGCFWMDNGHLTYREDQTSPPAGLRCLNMTDAQSGLASAP 60
 OY 61 VSGAGNHSYCRNPDEDPGRCVYSGAGVPEKRPCEDLRCPEETTSQALPAFTTEIOEAS 120
 DB 61 VSGAGNHSYCRNPDEDPGRCVYSGAGVPEKRPCEDLRCPEETTSQALPAFTTEIOEAS 120
 OY 121 EGPADDEVQVFAPANALPARSEAAAVQVIGISQVRMNSKEKDLGTGLGVLTVMVI 180
 DB 121 EGPADDEVQVFAPANALPARSEAAAVQVIGISQVRMNSKEKDLGTGLGVLTVMVI 180
 OY 181 IIAIGAGIILIGYSYKRGKDLKEQHDQVCEREMORTTLPLSAFNPTCEIVDEKTVVHT 240
 DB 181 IIAIGAGIILIGYSYKRGKDLKEQHDQVCEREMORTTLPLSAFNPTCEIVDEKTVVHT 240
 OY 241 SQTVPDPQEGSTPLMGQAGTPGA 263
 DB 241 SQTVPDPQEGSTPLMGQAGTPGA 263

RESULT 4

ID AAU86149
 AAU86149 standard; Protein: 263 AA.

XX AC AAU86149;
 XX DT 15-JUL-2002 (first entry)
 XX DE Human PRO264 polypeptide.
 XX KW Human; PRO, benign tumour; malignant tumour; lymphoid malignancy;
 KW leukaemia; neuronal disorder; stromal disorder; blastocytic disorder;
 KW inflammatory disorder; immune disorder; angiogenic disorder;
 KW cytoskeletal; neuroprotective.
 XX OS Homo sapiens.
 XX PN WO200153486-A1.
 XX PD 26-JUL-2001.
 XX PF 11-FEB-2000; 2000WO-US03565.
 XX PR 08-MAR-1999; 99MO-US050328.
 PR 11-MAR-1999; 99US-123972P.
 PR 11-MAY-1999; 99US-133459P.
 PR 02-JUN-1999; 99MO-US12252.
 PR 22-JUN-1999; 99US-140650P.
 PR 22-JUN-1999; 98US-140653P.
 PR 20-JUL-1999; 98US-144758P.
 PR 26-JUL-1999; 98US-145698P.
 PR 28-JUL-1999; 98US-146222P.
 PR 17-AUG-1999; 99US-149395P.
 PR 31-AUG-1999; 99US-151689P.

PR 01-SEP-1999; 99MO-US20111.
 PR 15-SEP-1999; 99MO-US21090.
 PR 30-NOV-1999; 99MO-US28311.
 PR 01-DEC-1999; 99MO-US28301.
 PR 01-DEC-1999; 99MO-US28634.
 PR 05-JAN-2000; 2000MO-US00219.
 XX (GENTH) GENENTECH INC.

XX Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
 PI Mazetars SA, Pan J, Plichi RM, Roy MA, Smith V, Stone DM;
 PI Matanabe CK, Wood WI;
 DR WPI, 2002-205567/26.
 DR N-PSDB; ABK40275.

XX Thirteen five nucleic acids encoding PRO polypeptides, useful for
 PT treating benign or malignant tumours, leukaemias and lymphoid
 PT malignancies, inflammatory, angiogenic and immunologic disorders -
 XX Claim 61; Fig 44; 302pp; English.

XX The present invention relates to the isolation of novel human PRO
 CC polypeptides and the polynucleotide sequences encoding them. The
 CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are
 CC useful for treating benign or malignant tumours (e.g. renal, kidney,
 CC bladder, breast, etc), leukaemias and lymphoid malignancies, other
 CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,
 CC macrophagal, stromal and blastocytic disorders, inflammatory, immune
 CC and angiogenic disorders. The polynucleotide sequences are also
 CC useful in gene therapy. AAU86128-AAU86162 represent the human PRO
 CC polypeptides of the invention.

XX Sequence 263 AA:

Query Match 99.8%; Score 1390; DB 23; Length 263;
 Best Local Similarity 99.8%; Pred. No. 1.7e-119;
 Matches 262; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLAMVQAFIVSNMLIAEYSGGCFWMDNGHLTYREDQTSPPAGLRCLNMTDAQSGLASAP 60
 DB 1 MLAMVQAFIVSNMLIAEYSGGCFWMDNGHLTYREDQTSPPAGLRCLNMTDAQSGLASAP 60
 OY 61 VSGAGNHSYCRNPDEDPGRCVYSGAGVPEKRPCEDLRCPEETTSQALPAFTTEIOEAS 120
 DB 61 VSGAGNHSYCRNPDEDPGRCVYSGAGVPEKRPCEDLRCPEETTSQALPAFTTEIOEAS 120
 OY 121 EGPADDEVQVFAPANALPARSEAAAVQVIGISQVRMNSKEKDLGTGLGVLTVMVI 180
 DB 121 EGPADDEVQVFAPANALPARSEAAAVQVIGISQVRMNSKEKDLGTGLGVLTVMVI 180
 OY 181 IIAIGAGIILIGYSYKRGKDLKEQHDQVCEREMORTTLPLSAFNPTCEIVDEKTVVHT 240
 DB 181 IIAIGAGIILIGYSYKRGKDLKEQHDQVCEREMORTTLPLSAFNPTCEIVDEKTVVHT 240
 OY 241 SQTVPDPQEGSTPLMGQAGTPGA 263
 DB 241 SQTVPDPQEGSTPLMGQAGTPGA 263

RESULT 5

ID AAB43237
 AAB43237 standard; Protein: 263 AA.

XX AC AAB43237;
 XX DT 08-FEB-2001 (first entry)
 XX DE Human ORFX ORF3001 polypeptide sequence SEQ ID NO:6002.
 XX KW Human; open reading frame; ORFX; detection; cytoskeletal; hepatotropic;
 KW vulnery; antiproliferative; antiparkinsonian; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;

immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 hypotensive; dermatological; immunosuppressive; antiinflammatory;
 antiviral; antibacterial; antifungal; antineumatic; antihypoid;
 antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 neurodegenerative disorder; osteoarthritis; graft vs host disease;
 cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 cholesterol ester storage; systemic lupus erythematosus; infection;
 severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 bone damage; cartilage damage; antiinflammatory disease; coagulation;
 thrombosis; contraceptive.
 Homo sapiens.
 OS
 XX
 PN WO2000058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000NO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-1) CURAGEN CORP.
 XX
 PI Shinketsu RA, Leach M;
 XX
 DR WPI; 2000-602362/57.
 DR N-PSDB; AAC77446.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 11; Page 5181-5182; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORF open reading frames 1 to 3161. The ORF
 CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
 CC antiproliferative; antiparkinsonian; nootropic; neuroprotective;
 CC osteoplastic; anticonvulsant; antitachytic; immunosuppressive;
 CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;
 CC antihypoid; and antianaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORF-associated disorder. The
 CC nucleic acids can be used to express ORF proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease, to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SO Sequence 263 AA:
 Query Match 99.4%; Score 1384; DB 21; Length 263;
 Best Local Similarity 99.2%; Pred. No. 66-119; Indels 0; Gaps 0;
 Matches 261; Conservative 1; Mismatches 1;
 QY 1 MLTAAVQAFIVSNMLAEAYGGGCFWNGHLYREDOTSPAPGLCLAWLDAOSGLASAP 60
 DB 1 MLTAAVQAFIVSNMLAEAYGGGCFWNGHLYREDOTSPAPGLCLAWLDAOSGLASAP 60
 QY 61 VSGAGNHSYCKNPDEPRGKWCYVSGEAGVPERKPCEDLRCEPTTSQALPAFTTEIOEAS 120
 DB 61 VSGAGNHSYCKNPDEPRGKWCYVSGEAGVPERKPCEDLRCEPTTSQALPAFTTEIOEAS 120

QY 121 EGRGADEVQVAPANALPARSEAAAVDPVIGISQVRNMSKEKKDLGTLGVYLGITMMVI 180
 DB 121 EGRGADEVQVAPANALPARSEAAAVDPVIGISQVRNMSKEKKDLGTLGVYLGITMMVI 180
 QY 181 IIAIAGGIIILGYSYKRGKDLKEHQDKVCEREMORITLPSAFTNPCEIYDEKTVVHT 240
 DB 181 IIAIAGGIIILGYSYKRGKDLKEHQDKVCEREMORITLPSAFTNPCEIYDEKTVVHT 240
 QY 241 SCPPVDPOEGSTPLMGAGCPGA 263
 DB 241 SCPPVDPOEGSTPLMGAGCPGA 263
 RESULT 6
 AAM93748
 ID AAM93748 standard; Protein; 263 AA.
 AC AAM93748;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide, SEQ ID NO: 3727.
 XX
 KW Human; full length cDNA; cDNA synthesis; oligo-capping.
 XX
 OS Homo sapiens.
 XX
 PN EP1130094-A2.
 XX
 PD 05-SEP-2001.
 XX
 PF 07-JUL-2000; 2000EP-0114089.
 XX
 PR 08-JUL-1999; 99JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Oka T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX
 DR WPI; 2001-524255/58.
 DR N-PSDB; AAK94700.
 XX
 PT 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -
 XX
 PS Claim 8; SEQ ID NO 3727; 1380pp + sequence listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 XX
 SO Sequence 263 AA:
 Query Match 99.0%; Score 1379; DB 22; Length 263;
 Best Local Similarity 99.2%; Pred. No. 1.7e-116; Indels 0; Gaps 0;
 Matches 261; Conservative 0; Mismatches 2;
 QY 1 MLTAAVQAFIVSNMLAEAYGGGCFWNGHLYREDOTSPAPGLCLAWLDAOSGLASAP 60
 DB 1 MLTAAVQAFIVSNMLAEAYGGGCFWNGHLYREDOTSPAPGLCLAWLDAOSGLASAP 60

| | | | |
|----------|---|---|-----|
| Qy | 61 | VSGAAGNSYCRNPDEDPEGPMPCVYSGAGAPKRCPCEDLCPPTTSQALPAFTTELQKAS | 120 |
| Qy | 61 | VSGAAGNSYCRNPDEDPEGPMPCVYSGAGAPKRCPCEDLCPPTTSQALPAFTTELQKAS | 120 |
| Db | 61 | VSGAAGNSYCRNPDEDPEGPMPCVYSGAGAPKRCPCEDLCPPTTSQALPAFTTELQKAS | 120 |
| Qy | 121 | BERGADGEVQFAPANALPAREEAAAVOPVYIGISGVPMNSKEKDDGLTGLVGLITMMVY | 180 |
| Db | 121 | BERGADGEVQFAPANALPAREEAAAVOPVYIGISGVPMNSKEKDDGLTGLVGLITMMVY | 180 |
| Qy | 181 | IIAIAAGIIGIGSYRKRGKDLKEQHDONCEREMOITLPISATNPFGSEYDEKTVYVHT | 240 |
| Db | 181 | IIAIAAGIIGIGSYRKRGKDLKEQHDONCEREMOITLPISATNPFGSEYDEKTVYVHT | 240 |
| Qy | 241 | SQTPVDPEGSSTPLMGOAGTGA 263 | |
| Db | 241 | SQTPVDPEGSSTPLMGOAGTGA 263 | |
| RESULT 7 | | | |
| ID | AA050220 | standard; Protein; 286 AA. | |
| AC | AA050220; | | |
| DT | 17-JUN-1999 | (first entry) | |
| XX | | Kringlet protein sequence. | |
| DE | | | |
| KV | | Kringlet; cancer; inflammation; autoimmunity; rheumatoid arthritis; AIDS; | |
| KM | | Parkinson's disease; cerebellar degeneration; Alzheimer's disease; asthma; | |
| KM | | Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis; | |
| KM | | neurological abnormalities; ischaemia reperfusion injury; ischaemic injury; | |
| KM | | cardiovascular disease; kidney disease; liver disease; aplastic anaemia; | |
| KM | | myocardial infarction; hypotension; hypertension; allergy; infection; | |
| KM | | myelodysplastic syndrome; haematologic abnormality; diagnosis; therapy; | |
| KM | | male pattern baldness. | |
| XX | | | |
| OS | Homo sapiens. | | |
| PN | W09911788-A1. | | |
| PD | 11-MAR-1999. | | |
| PF | 02-SEP-1998; | 98MO-US18270. | |
| PR | 01-SEP-1998; | 98US-0144889. | |
| PR | 02-SEP-1997; | 97US-0056032. | |
| XX | | (SMIK) SMITHKLINE BEECHAM CORP. | |
| XX | | | |
| XX | Albone EF, Kirkly KK; | | |
| DR | WPI, 1999-214707/18. | | |
| DR | N-PSDB; AAX28355. | | |
| PJ | | | |
| XX | | New kringlet polypeptides and polynucleotides | |
| PS | Claim 14; Page 33; 42p; English. | | |
| XX | | | |
| CC | This sequence is a Kringlet polypeptide of the invention. | | |
| CC | The kringlet polypeptides (I) are used to screen for agonists and | | |
| CC | antagonists. Agonists are used to treat subjects in need of enhanced | | |
| CC | activity or expression of (I). Antagonists are used to treat subjects | | |
| CC | having need to inhibit the activity or expression of (I). The methods can | | |
| CC | be used to treat conditions such as cancer, inflammation, autoimmunity, | | |
| CC | allergy, asthma, rheumatoid arthritis, CNS inflammation, cerebellar | | |
| CC | degeneration, Alzheimer's disease, Parkinson's disease, multiple | | |
| CC | sclerosis, amyotrophic lateral sclerosis, head injury damage and other | | |
| CC | neurological abnormalities, ischaemia reperfusion injury, cardiovascular | | |
| CC | disease, kidney disease, liver disease, ischaemic injury, myocardial | | |
| CC | infarction, hypotension, hypertension, AIDS, myelodysplastic syndromes | | |
| CC | and other haematologic abnormalities, aplastic anaemia, male pattern | | |
| CC | baldness, and bacterial, fungal, protozoan and viral infections. The | | |

| | |
|-----------------------|---|
| CC | Kringlet polypeptides may also be used to generate antibodies. |
| CC | Determining the presence or absence of mutations in, and analyzing for |
| CC | the presence or absence of expression of, kringlet polynucleotides can be |
| CC | used to diagnose a disease or susceptibility to a disease related to |
| CC | expression or activity of kringlet proteins. The polynucleotides may also |
| CC | be used for chromosome identification, and mapping. |
| XX | |
| SQ | Sequence 286 AA: |
| Query Match | 83.8% Score 1167.5; DB 20; Length 286; |
| Best Local Similarity | 87.3%; Pred. No. 5 2e-99; |
| Matches 227; | Conservative 5; Mismatches 27; Indels 1; Gaps 1 |
| Oy | 1 MLTAAVAFIVSNMILIANVGSGGCTNNGLTYEOTSPAGCPLCNMLDASGSLAP 60 |
| Db | 1 MLTAAVAFIVSNMILIANVGSGGCTNNGLTYEOTSPAGCPLCNMLDASGSLAP 60 |
| Oy | 61 VSGAGNNSYCRNPDEDPRGWCVSGSEAGVEPKPCEDLRCPETTSQALPAF-TTEIDEA 119 |
| Db | 61 VSGADNNSYCRNPDEDPRGWCVSGSEAGVEPKPCEDLRCPETTSQALPAFDNRSKR 120 |
| Oy | 120 SEGPDADEVVPANPALPARSBAAPVPIGISQRVRNNSKEKKDLGTLYVLGITMV 179 |
| Db | 121 LEPGADEVVVFPPANLPARSBALDPVIGISQRVDLGEKGGNSGLRAGHYHG 180 |
| Oy | 180 ILLAIGAGITLGYSYRKGDIKKHOKVCEREMORTPLPSAFTNPCTEIVDEKTYYH 239 |
| Db | 181 DIIAIGAGITLGYSYRKGDIKKHOKVCEREMORTPLPFAFTNPCEIIVDEKTYYH 240 |
| Oy | 240 TSGTPVDPQEGSTPLMGQG 259 |
| Db | 241 TSGTPVDPQEGSTPLMGQG 260 |
| RESULT 8 | |
| ID | ABB37905 |
| ID | ABB37905 standard; Peptide; 66 AA. |
| XX | |
| XX | ABB37905: |
| DT | 04-FEB-2002 (first entry) |
| XX | |
| DE | Peptide #5411 encoded by human foetal liver single exon probe. |
| XX | |
| KX | Human; foetal liver; gene expression; single exon nucleic acid probe. |
| OS | Homo sapiens. |
| XX | |
| FN | WC0200157277-A2. |
| PD | 09-AUG-2001. |
| PF | 30-JUN-2001; 2001WO-US00669. |
| XX | |
| PR | 04-FEB-2000; 2000US-0180312. |
| PR | 26-MAY-2000; 2000US-0207456. |
| PR | 30-JUN-2000; 2000US-0608408. |
| PR | 03-AUG-2000; 2000US-0632366. |
| PR | 21-SEP-2000; 2000US-0234687. |
| PR | 27-SEP-2000; 2000US-0236359. |
| PR | 04-OCT-2000; 2000GB-0024263. |
| XX | |
| PA | (MOLE-) MOLECULAR DYNAMICS INC. |
| PI | Penn SG, Hanzel DK, Chen W, Rank DR, |
| XX | |
| WT | WPI; 2001-483447/52. |
| XX | |
| PT | Human genome-derived single exon nucleic acid probes useful for |
| XX | analyzing gene expression in human fetal liver - |
| XX | |
| CS | Claim 27; SEQ ID NO 30540; 639bp + sequence listing; English. |

CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: the sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX

Sequence 66 AA;

Query Match 23.1%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 3.7e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 104 TTSQLPAPFTTEIOASESGPADEVGFAPANALPARESEAAVQPVIGISQRYVMNSKEK 163
DB 1 TTSQLPAPFTTEIOASESGPADEVGFAPANALPARESEAAVQPVIGISQRYVMNSKEK 60
|||
|||
|||

OY 164 KDLGTL 169
DB 61 KDLGTL 66
|||
|||
|||

RESULT 9
ID ABB23159 standard; Protein; 66 AA.
XX
AC ABB23159;
XX

DT 23-JAN-2002 (first entry)
XX

DE Protein #5158 encoded by probe for measuring heart cell gene expression.
XX

KM Human; gene expression; heart; microarray; vascular system;
KM cardiovascular disease; hypertension; cardiac arrhythmia;
KM congenital heart disease.
XX

OS Homo sapiens.
XX

PN WO200157274-A2.
XX

PD 09-AUG-2001.
XX

PF 30-JAN-2001; 2001WO-US00666.
XX

PR 04-FEB-2000; 2000US-0180312.
XX

PR 26-MAY-2000; 2000US-0207456.
XX

PR 30-JUN-2000; 2000US-0608408.
XX

PR 03-AUG-2000; 2000US-0632366.
XX

PR 21-SEP-2000; 2000US-0234687.
XX

PR 27-SEP-2000; 2000US-0236359.
XX

PR 04-OCT-2000; 2000GB-0024263.
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.
XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX

PT WPI; 2001-488899/53.
XX

PS Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX

CC Claim 15; SEQ ID NO 24929; 530bp; English.
XX

CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the

CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

Sequence 66 AA;

Query Match 23.1%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 3.7e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 104 TTSQLPAPFTTEIOASESGPADEVGFAPANALPARESEAAVQPVIGISQRYVMNSKEK 163
DB 1 TTSQLPAPFTTEIOASESGPADEVGFAPANALPARESEAAVQPVIGISQRYVMNSKEK 60
|||
|||
|||

OY 164 KDLGTL 169
DB 61 KDLGTL 66
|||
|||
|||

RESULT 10
ID AAM58537 standard; Protein; 66 AA.
XX
AC AAM58537;
XX

DT 05-NOV-2001 (first entry)
XX

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30642.
XX

KM Human; brain expressed exon; gene expression analysis; probe;
KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KM epilepsy; cancer.
XX

OS Homo sapiens.
XX

PN WO200157275-A2.
XX

PD 09-AUG-2001.
XX

PF 30-JAN-2001; 2001WO-US00667.
XX

PR 04-FEB-2000; 2000US-0180312.
XX

PR 26-MAY-2000; 2000US-0207456.
XX

PR 30-JUN-2000; 2000US-0608408.
XX

PR 03-AUG-2000; 2000US-0632366.
XX

PR 21-SEP-2000; 2000US-0234687.
XX

PR 27-SEP-2000; 2000US-0236359.
XX

PR 04-OCT-2000; 2000GB-0024263.
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.
XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX

PT WPI; 2001-483446/52.
XX

PS Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX

CC Example 4; SEQ ID NO: 30642; 650bp + sequence listing; English.
XX

CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancer. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX

Sequence 66 AA;

Query Match 23.1%; Score 322; DB 22; Length 66;

Best Local Similarity 100.0%; Pred. No. 3,7e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 104 TTSQLPAFTTEIQEASSEGPGADDEVQVAPANALPARSEAAAVQVIGISQVRNNSKEK 163
Db 1 TTSQLPAFTTEIQEASSEGPGADDEVQVAPANALPARSEAAAVQVIGISQVRNNSKEK 60
Oy 164 KDIGTL 169
Db 61 KDIGTL 66

RESULT 11

AAM71037
ID AAM71037 standard; Protein; 66 AA.

XX AAM71037;

XX 06-NOV-2001 (first entry)

XX Human bone marrow expressed probe encoded protein SEQ ID NO: 31343.

XX Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukemia; lymphoma; myeloma.

XX Homo sapiens.

XX W0200157276-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00668.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 31343; 658bp + Sequence listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX bone marrow. They can be used to measure gene expression in bone marrow

XX samples, which may enable the improved diagnosis and treatment of cancers

XX such as lymphoma, leukemia and myeloma. The present sequence is a

XX protein encoded by one of the probes of the invention.

XX Sequence 66 AA;

XX Query Match 23.1%; Score 322; DB 22; Length 66;

XX Best Local Similarity 100.0%; Pred. No. 3,7e-22;

XX Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 104 TTSQLPAFTTEIQEASSEGPGADDEVQVAPANALPARSEAAAVQVIGISQVRNNSKEK 163

Db 1 TTSQLPAFTTEIQEASSEGPGADDEVQVAPANALPARSEAAAVQVIGISQVRNNSKEK 60

Oy 164 KDIGTL 169

Db 61 KDIGTL 66

RESULT 12

AAM18800
ID AAM18800 standard; Protein; 66 AA.

XX AAM18800;

XX 12-OCT-2001 (first entry)

XX Peptide #5234 encoded by probe for measuring cervical gene expression.

XX Probe; human; microarray; gene expression; cervical epithelial cell;

XX cervical cancer.

XX Homo sapiens.

XX W0200157278-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00670.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488901/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human cervical epithelial cells -

XX Claim 27; SEQ ID No 23626; 487bp; English.

XX The present invention relates to human single exon nucleic acid probes

XX (SENP: see A110068-A1128459). The present sequence is a peptide encoded

XX by one such probe. The SENPs are derived from human HeLa cells. The SENPs

XX can be used to produce a single exon microarray, which can be used for

XX measuring human gene expression in a sample derived from human cervical

XX epithelial cells. By measuring gene expression, the probes are therefore

XX useful in grading and/or staging of diseases of the cervix, notably

XX cervical cancer.

XX Note: The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at http://wipo.int/pub/published_pct_sequences.

XX Sequence 66 AA;

XX Query Match 23.1%; Score 322; DB 22; Length 66;

XX Best Local Similarity 100.0%; Pred. No. 3,7e-22;

XX Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 104 TTSQLPAFTTEIQEASSEGPGADDEVQVAPANALPARSEAAAVQVIGISQVRNNSKEK 163

Db 1 TTSQLPAFTTEIQEASSEGPGADDEVQVAPANALPARSEAAAVQVIGISQVRNNSKEK 60

Oy 164 KDIGTL 169

Db 61 KDIGTL 66

RESULT 13

AAM31314

ID AAM31314 standard; Protein; 66 AA.

XX AAM31314;

XX 17-OCT-2001 (first entry)
XX Peptide #5351 encoded by probe for measuring placental gene expression.
DE Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
XX Homo sapiens.
XX WO200157272-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00663.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX Claim 27; SEQ ID No 31583; 654pp; English.
XX The present invention relates to single exon nucleic acid probes (SENP:
XX see A1131315-A157546). The present sequence is a peptide encoded by one
XX such probe. The probes are useful for producing a microarray for
XX predicting, measuring and displaying gene expression in samples derived
XX from human placenta. The probes are useful for antenatal diagnosis of
XX human genetic disorders.
XX Sequence 66 AA;
SQ Query Match 23.1%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 3.7e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 104 TTGQALPAFTTETQASSEGGADEVGFAPANALPARSEAAVCPVIGISORVRNNSKEK 163
DB 1 TTGQALPAFTTETQASSEGGADEVGFAPANALPARSEAAVCPVIGISORVRNNSKEK 60
QY 164 KDLGTL 169
DB 61 KDLGTL 66
RESULT 14
ABG40828
ID ABG40828 standard; Peptide; 66 AA.
XX ABG40828;
XX 19-AUG-2002 (first entry)
XX Human peptide encoded by genome-derived single exon probe SEQ ID 30493.
XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
XX Chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Kartagener syndrome;

KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX Homo sapiens.
XX WO200186003-A2.
XX 15-NOV-2001.
XX 30-JAN-2001; 2001WO-US00665.
XX 04-FEB-2000; 2000US-180312P.
XX 26-MAY-2000; 2000US-207456P.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-234687P.
XX 27-SEP-2000; 2000US-236359P.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples -
XX Claim 27; SEQ ID No 30493; 634pp; English.
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of
XX probes; the novel set of probes which hybridise at high stringency to a
XX nucleic acid expressed in the human lung; measuring gene expression in a
XX sample derived from human lung, comprising (a) contacting the array with
XX a collection of detectably labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the label detectably bound to each probe of
XX the array; identifying exons in a eukaryotic genome, comprising
XX (a) algorithmically predicting at least one exon from genomic sequences
XX of the eukaryote; and (b) detecting specific hybridisation of detectably
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon, the probe is included
XX in the above mentioned microarray; assigning exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridisation to a single exon
XX microarray having a probe with the exon, where a common pattern of
XX expression of the exons in the tissues and/or cell types indicates that
XX the exons should be assigned to a single gene; a peptide comprising one
XX of 12011 sequences, mentioned in the specification, or encoded by the
XX probes/open reading frames (ORF). The probes are used for gene
XX expression analysis, and for identifying exons in a gene, particularly
XX using human lung derived mRNA and for the study of lung diseases
XX such as asthma, lung cancer, chronic obstructive pulmonary disease
XX (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
XX fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
XX Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
XX haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
XX pulmonary alveolar proteinosis, Kartagener syndrome, fibrocystic
XX pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
XX and hyaline membrane disease. The present sequence is a peptide/protein
XX encoded by a single exon probe of the invention.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp.wipo.int/pub/published_pcl_sequences.
SQ Sequence 66 AA;

Query Match 23.1%; Score 322; DB 23; Length 66;
 Best Local Similarity 100.0%; Pred. No. 3,7e-22;
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 104 TTSGALPAFTTEIOEASGEGADEVOVFAPANALPARSAAAVPVIGISQRYVMSKEK 163
 Db 1 TTSGALPAFTTEIOEASGEGADEVOVFAPANALPARSAAAVPVIGISQRYVMSKEK 60

Oy 164 KDLCGL 169
 Db 61 KDLCGL 66

RESULT 15

AA12615
 ID AA12615 standard; Protein; 56 AA.

XX AC AA12615;

XX DT 22-JUN-1999 (first entry)

XX DE Human 5' EST secreted protein SEQ ID NO: 280 from WO 9906553.

XX KM Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KM forensic; gene therapy; chromosome mapping; signal peptide;
 KM upstream regulatory sequence; cytokine activity; cell proliferation;
 KM differentiation; haematopoiesis regulation; tissue growth regulation;
 KM reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 KM thrombolytic; antiinflammatory; tumour inhibition; antitumour.

XX OS Homo sapiens.

XX FN WO9906553-A2.

XX PD 11-FEB-1999.

XX PF 31-JUL-1998; 98WO-IB01237.

XX PR 01-AUG-1997; 97US-0905051.

XX PA (GEST) GENSET.

XX PI Duclert A, Dumas Milne Edwards J, Lacroix B;

XX DR WP1; 1999-153783/13.

XX DR N-PSDB; AA141473.

XX PT New nucleic acids encoding human secreted proteins - obtained from
 XX CDNA libraries derived from umbilical cord, lymph ganglia,
 XX lymphocytes and placental tissue

XX PS Claim 34; Page 376; 411pp; English.

XX AA141379 to AA141526 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AA12521 to
 CC AA12668, respectively. The proteins given represent the signal peptide
 CC and an N-terminal fragment of a secreted protein. The nucleic acid
 CC sequences can be used for producing secreted human gene products. They
 CC can also be used to develop products for diagnosis and therapy. The
 CC proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, haematopoiesis regulating
 CC activity, tissue growth regulating activity, reproductive hormone
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
 CC thrombolytic activity, receptor/ligand activity, antiinflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.

XX Sequence 56 AA;

Query Match 22.0%; Score 306; DB 20; Length 56;
 Best Local Similarity 98.2%; Pred. No. 8.8e-21;
 Matches 55; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MLLAWOAFIVSNMLAEAYSGGCFMDNGHLTYREDQTSAPAGLRCLNWLDAQSGL 56
 Db 1 MLLAWOAFIVSNMLAEAYSGGCFMDNGHLTYREDQTSAPAGLRCLNWLDAQSGL 56

Search completed: April 7, 2003, 06:59:47
 Job time : 43.3589 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 06:48:07 ; Search time 14.8468 Seconds

(without alignments)
1082.981 Million cell updates/sec

Title: US-10-057-951-2

Perfect score: 1393

Sequence: 1 MLAWVOAFLVSNMLAEAY.....PVDPOGSGPTLMGQGTGCA 263

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PC7_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/PCUTS_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------------|---------------------|
| 1 | 1393 | 100.0 | 263 | 9 US-10-102-704-2 | Sequence 2, Appl1 |
| 2 | 1393 | 100.0 | 263 | 10 US-10-057-951-2 | Sequence 2, Appl1 |
| 3 | 1393 | 100.0 | 263 | 10 US-09-084-491A-2 | Sequence 2, Appl1 |
| 4 | 1322 | 23.1 | 527 | 9 US-09-864-457-18 | Sequence 38457, A |
| 5 | 154.5 | 11.1 | 527 | 9 US-09-987-455-18 | Sequence 18, Appl1 |
| 6 | 154.5 | 11.1 | 527 | 9 US-09-987-455-18 | Sequence 19, Appl1 |
| 7 | 154.5 | 11.1 | 527 | 9 US-09-987-455-18 | Sequence 145, Appl1 |
| 8 | 154.5 | 11.1 | 527 | 9 US-09-987-455-18 | Sequence 7, Appl1 |
| 9 | 142 | 10.2 | 160 | 9 US-09-969-271-7 | Sequence 35, Appl1 |
| 10 | 142 | 10.2 | 160 | 9 US-09-969-271-7 | Sequence 35, Appl1 |
| 11 | 142 | 10.2 | 160 | 9 US-09-969-271-7 | Sequence 35, Appl1 |
| 12 | 142 | 10.2 | 160 | 9 US-09-969-271-7 | Sequence 35, Appl1 |
| 13 | 142 | 10.2 | 250 | 9 US-09-335-325-30 | Sequence 30, Appl1 |
| 14 | 142 | 10.2 | 250 | 9 US-09-335-325-30 | Sequence 30, Appl1 |
| 15 | 142 | 10.2 | 250 | 9 US-09-335-325-30 | Sequence 30, Appl1 |
| 16 | 142 | 10.2 | 250 | 9 US-09-335-325-30 | Sequence 30, Appl1 |
| 17 | 142 | 10.2 | 339 | 9 US-09-335-325-30 | Sequence 3, Appl1 |
| 18 | 142 | 10.2 | 339 | 9 US-09-335-325-30 | Sequence 3, Appl1 |
| 19 | 142 | 10.2 | 339 | 10 US-09-761-120-3 | Sequence 3, Appl1 |

| | | | | | |
|----|-----|------|-----|---------------------|--------------------|
| 20 | 142 | 10.2 | 352 | 9 US-09-335-325-40 | Sequence 40, Appl1 |
| 21 | 142 | 10.2 | 352 | 9 US-10-111-241-40 | Sequence 40, Appl1 |
| 22 | 142 | 10.2 | 352 | 10 US-09-761-120-40 | Sequence 40, Appl1 |
| 23 | 142 | 10.2 | 368 | 10 US-09-761-120-42 | Sequence 42, Appl1 |
| 24 | 142 | 10.2 | 378 | 9 US-09-335-325-42 | Sequence 42, Appl1 |
| 25 | 142 | 10.2 | 378 | 9 US-10-111-241-42 | Sequence 42, Appl1 |
| 26 | 142 | 10.2 | 378 | 9 US-09-873-676-1 | Sequence 1, Appl1 |
| 27 | 142 | 10.2 | 458 | 10 US-09-946-893-4 | Sequence 4, Appl1 |
| 28 | 142 | 10.2 | 569 | 10 US-09-946-893-5 | Sequence 5, Appl1 |
| 29 | 142 | 10.2 | 571 | 10 US-09-946-893-8 | Sequence 8, Appl1 |
| 30 | 142 | 10.2 | 576 | 10 US-09-946-893-6 | Sequence 6, Appl1 |
| 31 | 142 | 10.2 | 791 | 9 US-09-967-186-1 | Sequence 1, Appl1 |
| 32 | 142 | 10.2 | 810 | 10 US-09-946-893-2 | Sequence 2, Appl1 |
| 33 | 138 | 9.9 | 79 | 9 US-09-335-325-5 | Sequence 8, Appl1 |
| 34 | 138 | 9.9 | 79 | 9 US-10-111-241-8 | Sequence 11, Appl1 |
| 35 | 138 | 9.9 | 79 | 9 US-09-753-064-2 | Sequence 8, Appl1 |
| 36 | 138 | 9.9 | 79 | 10 US-09-753-064-2 | Sequence 8, Appl1 |
| 37 | 135 | 9.7 | 377 | 9 US-09-987-455-8 | Sequence 8, Appl1 |
| 38 | 135 | 9.5 | 326 | 9 US-10-057-951-3 | Sequence 8, Appl1 |
| 39 | 133 | 9.5 | 354 | 9 US-09-987-457-10 | Sequence 3, Appl1 |
| 40 | 133 | 9.5 | 354 | 9 US-09-987-455-11 | Sequence 10, Appl1 |
| 41 | 133 | 9.5 | 372 | 9 US-10-102-704-3 | Sequence 3, Appl1 |
| 42 | 133 | 9.5 | 372 | 10 US-09-084-491A-3 | Sequence 3, Appl1 |
| 43 | 132 | 9.5 | 411 | 9 US-09-880-503-3 | Sequence 3, Appl1 |
| 44 | 132 | 9.5 | 431 | 9 US-10-076-421-2 | Sequence 2, Appl1 |
| 45 | 132 | 9.5 | 431 | 10 US-09-264-468B-1 | Sequence 1, Appl1 |

ALIGNMENTS

RESULT 1
US-10-102-704-2
; Sequence 2, Application US/10102704
; Patient No. US20020164768A1
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OR INVENTION: Tissue Plasminogen Activator-Like Protein
; FILE REFERENCE: P37378C1
; CURRENT APPLICATION NUMBER: US/10/102,704
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/084,491
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: 60/048,000
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-102-704-2

Query Match 100.0%; Score 1393; DB 9; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.6e+19;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-----|
| QY | 1 | MLAWVOAFLVSNMLAEAYGSGGCPMNGHLYREOTSPARGIACIWMLOAGLASAP | 60 |
| DB | 1 | MLAWVOAFLVSNMLAEAYGSGGCPMNGHLYREOTSPARGIACIWMLOAGLASAP | 60 |
| QY | 61 | VSGAONHSYCNPEDEPRGPGWCYSGEAGVPERKPCEDLRCPETTSQALPAFTTEIOEAS | 120 |
| DB | 61 | VSGAONHSYCNPEDEPRGPGWCYSGEAGVPERKPCEDLRCPETTSQALPAFTTEIOEAS | 120 |
| QY | 121 | EGPGADEVOVAPAPNALPARSBAAVQFVIGISQRYVMSKEKKDLGTLGVILGITMWT | 180 |
| DB | 121 | EGPGADEVOVAPAPNALPARSBAAVQFVIGISQRYVMSKEKKDLGTLGVILGITMWT | 180 |
| QY | 181 | IIAIGAGIILISYKRGDKKEQHQKQCEREMORITLPSAFTNPCEIIVDEKTVVHT | 240 |
| DB | 181 | IIAIGAGIILISYKRGDKKEQHQKQCEREMORITLPSAFTNPCEIIVDEKTVVHT | 240 |

Mon Apr 7 07:36:46 2003

us-10-057-951-2.rapb

Page 2

QY 241 SQTVPDPOEGSTPLMGAGTGA 263
Db 241 SQTVPDPOEGSTPLMGAGTGA 263

RESULT 2

US-10-057-951-2
Sequence 2, Application US/10057951
Patent No. US20020177213A1
GENERAL INFORMATION:
APPLICANT: MOORE et al.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
FILE REFERENCE: PF378P1
CURRENT APPLICATION NUMBER: US/10/057,951
PRIOR FILING DATE: 2002-01-29
PRIOR APPLICATION NUMBER: US 09/411,977
PRIOR FILING DATE: 1999-10-04
PRIOR APPLICATION NUMBER: US 09/084,491
PRIOR FILING DATE: 1998-05-27
PRIOR APPLICATION NUMBER: US 60/048,000
PRIOR FILING DATE: 1997-05-28
NUMBER OF SEQ ID NOS: 31
SOFTWARE: Patentin Ver. 3.1
SEQ ID NO 2
LENGTH: 263
TYPE: PRP
ORGANISM: Homo sapiens
US-10-057-951-2

Query Match 100.0%; Score 1393; DB 9; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.6e-119; Indels 0; Gaps 0;
Matches 263; Conservative 0; Mismatches 0;

QY 1 MLAWVOAFVLSNMLAEAYSGGCFMNDGHLVREDQTSPPAGRLCLNMLDAQSGLASAP 60
Db 1 MLAWVOAFVLSNMLAEAYSGGCFMNDGHLVREDQTSPPAGRLCLNMLDAQSGLASAP 60
QY 61 VSGAGNHSYCRNPDEDPFGPCVYSGAGVPEKRPCEDLRCPETTSQALPFTTEIOEAS 120
Db 61 VSGAGNHSYCRNPDEDPFGPCVYSGAGVPEKRPCEDLRCPETTSQALPFTTEIOEAS 120
QY 121 EGPQADEVOVFAFANALPARSEAAAVOPVIGISQVRNMSKEKDLGTLGYVIGITMVI 180
Db 121 EGPQADEVOVFAFANALPARSEAAAVOPVIGISQVRNMSKEKDLGTLGYVIGITMVI 180
QY 181 IIAIGAGIIGYSYKRGKDLKEQHDQKVCEREMQRTITPLSAFTNPCEIYDEKTVVHT 240
Db 181 IIAIGAGIIGYSYKRGKDLKEQHDQKVCEREMQRTITPLSAFTNPCEIYDEKTVVHT 240
QY 241 SQTVPDPOEGSTPLMGAGTGA 263
Db 241 SQTVPDPOEGSTPLMGAGTGA 263

RESULT 3

US-09-084-491A-2
Sequence 2, Application US/09084491A
Patent No. US20020061576A1
GENERAL INFORMATION:
APPLICANT: MOORE, PAUL A.
APPLICANT: RUBEN, STEVEN M.
APPLICANT: EBNER, REINHARD
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESS: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/084,491A
FILING DATE: 27-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PR378
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-084-491A-2

Query Match 100.0%; Score 1393; DB 10; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.6e-119; Indels 0; Gaps 0;
Matches 263; Conservative 0; Mismatches 0;

QY 1 MLAWVOAFVLSNMLAEAYSGGCFMNDGHLVREDQTSPPAGRLCLNMLDAQSGLASAP 60
Db 1 MLAWVOAFVLSNMLAEAYSGGCFMNDGHLVREDQTSPPAGRLCLNMLDAQSGLASAP 60
QY 61 VSGAGNHSYCRNPDEDPFGPCVYSGAGVPEKRPCEDLRCPETTSQALPFTTEIOEAS 120
Db 61 VSGAGNHSYCRNPDEDPFGPCVYSGAGVPEKRPCEDLRCPETTSQALPFTTEIOEAS 120
QY 121 EGPQADEVOVFAFANALPARSEAAAVOPVIGISQVRNMSKEKDLGTLGYVIGITMVI 180
Db 121 EGPQADEVOVFAFANALPARSEAAAVOPVIGISQVRNMSKEKDLGTLGYVIGITMVI 180
QY 181 IIAIGAGIIGYSYKRGKDLKEQHDQKVCEREMQRTITPLSAFTNPCEIYDEKTVVHT 240
Db 181 IIAIGAGIIGYSYKRGKDLKEQHDQKVCEREMQRTITPLSAFTNPCEIYDEKTVVHT 240
QY 241 SQTVPDPOEGSTPLMGAGTGA 263
Db 241 SQTVPDPOEGSTPLMGAGTGA 263

RESULT 4

US-09-864-761-38457
Sequence 38457, Application US/09864761
Patent No. US20020046763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30

```
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 38457
/ LENGTH: 66
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC002073.1
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
/ OTHER INFORMATION: EXPRESSED IN HEAL, SIGNAL = 2.3
/ OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 1.9
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
/ OTHER INFORMATION: SWISSPROT HIT: P39881, EVALU 3.40e+00
/ OTHER INFORMATION: EST_HUMAN HIT: BE299406.1, EVALU 5.00e-30
/ US-09-864-761-38457

Query Match      23.1%; Score 322; DB 10; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.6e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 104 TTSSALPATTETIOASGSGPADDEVUPAPANALPARSBAAYOVPGISQRYMNSKEX 163
Db 1 TTSSALPATTETIOASGSGPADDEVUPAPANALPARSBAAYOVPGISQRYMNSKEX 60

Qy 164 KDLGTL 169
Db 61 KDLGTL 66

RESULT 5
US-09-987-457-18
/ Sequence 18, Application US/09987457
/ Publication No. US20030013150A1
/ GENERAL INFORMATION:
/ APPLICANT: Manosroi, Aranya
/ APPLICANT: Manosroi, Jiradej
/ APPLICANT: Tayapiwatana, Chatchai
/ APPLICANT: Goetz, Friedrich
/ APPLICANT: Werner, Rolf-Guenther
/ TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
/ FILE REFERENCE: 0652.2180001
/ CURRENT APPLICATION NUMBER: US/09/987,457
/ CURRENT FILING DATE: 2001-11-14
/ PRIOR APPLICATION NUMBER: 60/268,573
/ PRIOR FILING DATE: 2001-02-15
```

```
/ PRIOR APPLICATION NUMBER: GB 00 27 782.2
/ PRIOR FILING DATE: 2000-11-14
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 18
/ LENGTH: 527
/ TYPE: PRT
/ ORGANISM: Homo sapiens (CPA)
/ US-09-987-457-18

Query Match      11.1%; Score 154.5; DB 9; Length 527;
Best Local Similarity 39.6%; Pred. No. 4.6e-06;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

Qy 25 CFMDNGHLYREDQTSFAPGLRCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 75
Db 92 CYEDOGISYRGTSVTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGIQHNNYCRNPDR 149

Qy 76 DPRGPMCYSGEAGVPEKRPCEDLRCPEPTTS 106
Db 150 DSK-PMCYVF-KAGKYSSEFCSTPACSEGNS 178

RESULT 6
US-09-987-455-19
/ Sequence 19, Application US/09987455
/ Publication No. US20030049729A1
/ GENERAL INFORMATION:
/ APPLICANT: Aranya, Manosroi
/ APPLICANT: Chatchai, Tayapiwatana
/ APPLICANT: Goetz, Friedrich
/ APPLICANT: Rolf-Guenther, Werner
/ TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
/ FILE REFERENCE: 0652.2190001
/ CURRENT APPLICATION NUMBER: US/09/987,455
/ CURRENT FILING DATE: 2001-11-14
/ PRIOR APPLICATION NUMBER: 60/268,574
/ PRIOR FILING DATE: 2001-02-15
/ PRIOR APPLICATION NUMBER: GB 0027779.8
/ PRIOR FILING DATE: 2000-11-14
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 19
/ LENGTH: 527
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-987-455-19

Query Match      11.1%; Score 154.5; DB 9; Length 527;
Best Local Similarity 39.6%; Pred. No. 4.6e-06;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

Qy 25 CFMDNGHLYREDQTSFAPGLRCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 75
Db 92 CYEDOGISYRGTSVTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGIQHNNYCRNPDR 149

Qy 76 DPRGPMCYSGEAGVPEKRPCEDLRCPEPTTS 106
Db 150 DSK-PMCYVF-KAGKYSSEFCSTPACSEGNS 178

RESULT 7
US-09-974-298-145
/ Sequence 145, Application US/09974298
/ Patent No. US20020156263A1
/ GENERAL INFORMATION:
/ APPLICANT: Chen, Huei-Mei
/ TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
/ FILE REFERENCE: PA-0037 P
/ CURRENT APPLICATION NUMBER: US/09/974,298
/ CURRENT FILING DATE: 2001-10-04
```



```

CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30303-1769

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/335,325
FILING DATE: 17-Jun-1999
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/612,788
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Warren, William L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 05213-0126
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 250 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: K1-3
SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-09-335-325-30

Query Match 10.2%; Score 142; DB 9; Length 250;
Best Local Similarity 37.2%; Pred. No. 2,4e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 25 CFWDNGHLYREDQTSPPAGRLCLMWLDA-----OSGLASAVSGAGNHSYCRNPDEDPBG 79
Db 1 CKTGNGKRYRGTMTSTKNGITCOKWSTSPRPRPTATPSEGL-EEHYCRNPUNDPQG 59
QY 80 PWCYVSGEAGVPEKR--PCEDLRCE 103
Db 60 PWCYTTD---PEKRYDYCDILECEE 81

RESULT 13
US-10-131-241-30
; Sequence 30, Application US/10131241
; Publication No. US20030012792A1
GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortlier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolife
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin version 3.1

```

```

; SEQ ID NO 30
; LENGTH: 250
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-131-241-30

```

| | | | | |
|--------------------------|--------|--------------------|------------|-------------|
| Query Match | 10.2%; | Score 142; | DB 9; | length 250; |
| Best Local Similarity | 37.2%; | Pred. No. 2.4e-05; | | |
| Matches 32; Conservative | 8; | Mismatches 34; | Indels 12; | Gaps 4; |

QY 25 CFFDNGHLRYREDQNSPAPGLRCLMLUDA-----QSGLASAPVSGAGNHSYCNAPDEDEPRG 75
 | | | | : | : | : | | | | | : |
 Db 1 CKTGNGKNYRGTMSTKNGITCQKMSSTSPHPRFESPATHPSEGL-ENNYCRNPNDNDFQG 55

```

QY      80  PWCYVSGEAGVPEKR--PCEDLRCP 103
          ||||: ||||| |:|||
Db      60  PWCYTTD---PEKRYDYCDILECEE 81

```

```

RESULT 14
US-09-761-120-30
Sequence 30, Application US/09761120
Patent No. US28020037847A1
GENERAL INFORMATION:
APPLICANT: Folkman, M, Judah
APPLICANT: O'Reilly, Michael
TITLE OF INVENTION: Nucleic Acids Encoding K-15 Region Fragments of Plasminogen
FILE REFERENCE: 05940-0151 (43171-252068)
CURRENT APPLICATION NUMBER: US/09/761,120
PRIORITY FILING DATE: 2001-01-16
PRIORITY APPLICATION NUMBER: 09/309,821
PRIORITY FILING DATE: 1999-05-11
PRIORITY APPLICATION NUMBER: 08/866,735
PRIORITY FILING DATE: 1997-05-30
NUMBER OF SEQ ID NOS: 47
SOFTWARE: Patentin version 3.0
SEQ ID NO 30
LENGTH: 250
TYPE: PR1
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: K-15 Region Fragments of Plasminogen
US-09-761-120-30

```

| | | | | |
|--------------------------|-------|--------------------|------------|-------------|
| Query Match | 10.2% | Score 142; | DB 10; | Length 250; |
| Best Local Similarity | 37.2% | Pred. No. 2.4e-05; | | |
| Matches 32; Conservative | 8; | Mismatches 34; | Indels 12; | Gaps 4 |

QY 25 CFWNGHLTYREDQTSBPAGRLCLNWLDA-----QSGLASAPVSGAGNHSYCRNPDEDEPRG 74

DB 1 CKTGNGKNTYRGTMSTKTKNGITPCQKMSSTSPHPRPRFSPTATHSPRL--EENYICAPDPNDPQG 5

```

QY      80  PWCYVSGEAGVPEKR--PCEDLRCP 10
          ||||: ||||: ||||
Db      60  PWCYTTD---PEKRYDCDILECEE 81

```

```

RESULT 15
US-10-131-241-61
; Sequence 61, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Portier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802

```

```

? PRIOR FILING DATE: 1999-05-21
? PRIOR APPLICATION NUMBER: US 60/086,586
? PRIOR FILING DATE: 1998-05-22
? NUMBER OF SEQ ID NOS: 65
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 61
? LENGTH: 260
? TYPE: PRT
? ORGANISM: Homo sapiens
US-10-131-241-61

```

| | | | | |
|-----------------------|--------------|--------------------|----------------|-------------|
| Query Match | 10.2% | Score 142; | DB 9; | length 260; |
| Best Local Similarity | 37.2%; | Pred. No. 2.5e-05; | | |
| Matches 32; | Conservative | 8; | Mismatches 34; | Indels 12; |
| | | | Gaps | 4 |

DQ 25 CFFNDNGHLYREDQTSAPQLRLWLDA-----QSGLASAPVSGAGNHSYCNPDDEBPRG 79
| | | : | : | : | : | : | : | :
DB 6 CKTGNCKNYRGITMSKTKNGITTCQKMSSTSPHRPRTSPATHPSEGL-EENYCNPDNDPQG 64

```

QY      80  PWCYVSGEAGVPEKR--PCEDLRCP 10
          ||| : ||| : |||
Db      65  PWCYTTD---PEKRYDYCDILECEE 86

```

Search completed: April 7, 2003, 07:01:32
Job time : 15.8468 secs

Search completed: April 7, 2003, 07:01:32
Job time : 15.8468 secs

GenCore version 5.1.4.B5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 06:37:51 ; Search time 192.301 Seconds

(without alignments)
881.767 Million cell updates/sec

Title: US-10-057-951-2

Perfect score: 1393
Sequence: 1 MLAWVQAFVLSNMILAEAY.....FVDPQSGSTPLMGQAGTPGA 263

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*

2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*

3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*

4: /cgn2_6/ptodata/1/paa/US08_COMB.pep:*

5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*

6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*

7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*

8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*

9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*

10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*

11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*

12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*

13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*

14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*

15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*

16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*

17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*

18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*

19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*

20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*

21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*

22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*

23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*

24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*

25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*

26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*

27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 1393 | 100.0 | 263 | 1 | PCT-US00-27239-2 |
| 2 | 1393 | 100.0 | 263 | 1 | PCT-US02-05301-164 |
| 3 | 1393 | 100.0 | 263 | 1 | PCT-US98-10728-2 |
| 4 | 1393 | 100.0 | 263 | 14 | US-09-084-491A-2 |
| 5 | 1393 | 100.0 | 263 | 15 | US-09-144-889-2 |
| 6 | 1393 | 100.0 | 263 | 24 | US-10-057-951-2 |

| | | | | | | |
|----|--------|-------|-----|----|----------------------|--------------------|
| 7 | 1393 | 100.0 | 263 | 25 | US-10-102-704-2 | Sequence 2, Appl1 |
| 8 | 1393 | 100.0 | 263 | 27 | US-60-048-000-2 | Sequence 2, Appl1 |
| 9 | 1390 | 99.8 | 263 | 21 | US-09-791-537-39562 | Sequence 39562, A |
| 10 | 1390 | 99.8 | 263 | 23 | US-09-927-796-44 | Sequence 44, Appl1 |
| 11 | 1390 | 99.8 | 263 | 26 | US-10-210-951-44 | Sequence 44, Appl1 |
| 12 | 1390 | 99.8 | 263 | 26 | US-10-211-858-44 | Sequence 44, Appl1 |
| 13 | 1390 | 99.8 | 263 | 26 | US-10-211-884-44 | Sequence 44, Appl1 |
| 14 | 1390 | 99.8 | 263 | 27 | US-60-230-435-1580 | Sequence 1580, Ap |
| 15 | 1390 | 99.8 | 263 | 27 | US-60-389-987-244 | Sequence 244, App |
| 16 | 1390 | 99.8 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 17 | 1379 | 99.0 | 263 | 20 | US-09-611-526-3727 | Sequence 3727, Ap |
| 18 | 1339 | 96.1 | 257 | 1 | PCT-US01-14827-8858 | Sequence 8858, Ap |
| 19 | 1328 | 95.3 | 250 | 14 | US-09-001-403-64 | Sequence 64, Appl1 |
| 20 | 1325.5 | 95.2 | 295 | 1 | PCT-US01-14827-8900 | Sequence 8900, Ap |
| 21 | 1270.5 | 91.2 | 308 | 27 | US-60-197-873-13839 | Sequence 471, App |
| 22 | 1167.5 | 83.8 | 286 | 15 | US-09-144-889-4 | Sequence 4, Appl1 |
| 23 | 953 | 68.4 | 178 | 27 | US-60-213-800-237 | Sequence 237, App |
| 24 | 752 | 54.0 | 146 | 1 | PCT-US02-05301-311 | Sequence 311, App |
| 25 | 752 | 54.0 | 146 | 1 | PCT-US02-05301-311 | Sequence 311, App |
| 26 | 700 | 50.3 | 126 | 17 | US-09-307-140-979 | Sequence 979, App |
| 27 | 700 | 50.3 | 126 | 22 | US-09-817-076-979 | Sequence 979, App |
| 28 | 612 | 43.9 | 109 | 22 | US-09-834-366-13839 | Sequence 13839, A |
| 29 | 612 | 43.9 | 109 | 27 | US-60-197-873-13839 | Sequence 13839, A |
| 30 | 580 | 41.6 | 117 | 1 | PCT-US01-14827-8897 | Sequence 8897, Ap |
| 31 | 426 | 30.6 | 88 | 22 | US-09-834-366-16200 | Sequence 16200, A |
| 32 | 426 | 30.6 | 88 | 27 | US-60-197-873-16200 | Sequence 16200, A |
| 33 | 322 | 23.1 | 66 | 1 | PCT-US01-00663-31583 | Sequence 31583, A |
| 34 | 322 | 23.1 | 66 | 22 | US-09-864-761-38457 | Sequence 38457, A |
| 35 | 322 | 23.1 | 66 | 25 | US-10-182-993-30642 | Sequence 30642, A |
| 36 | 322 | 23.1 | 66 | 25 | US-10-182-993-24929 | Sequence 24929, A |
| 37 | 322 | 23.1 | 66 | 26 | US-10-182-997-23626 | Sequence 23626, A |
| 38 | 322 | 23.1 | 66 | 26 | US-10-203-134-31343 | Sequence 31343, A |
| 39 | 322 | 23.1 | 66 | 26 | US-10-203-135-30493 | Sequence 30493, A |
| 40 | 322 | 23.1 | 66 | 26 | US-10-203-136-31400 | Sequence 31400, A |
| 41 | 322 | 23.1 | 66 | 26 | US-10-203-137-31583 | Sequence 31583, A |
| 42 | 322 | 23.1 | 66 | 26 | US-10-203-139-30540 | Sequence 30540, A |
| 43 | 306 | 22.0 | 56 | 13 | US-08-905-051-280 | Sequence 280, App |
| 44 | 306 | 22.0 | 56 | 13 | US-08-947-599C-2338 | Sequence 2338, App |
| 45 | 306 | 22.0 | 56 | 13 | US-08-905-135-428 | Sequence 428, App |

ALIGNMENTS

RESULT 1

PCT-US00-27239-2

; Sequence 2, Application PC/TUS0027239

; GENERAL INFORMATION:

; APPLICANT: HUMAN GENOME SCIENCES, INC.

; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease

; FILE REFERENCE: PF378PCT2

; CURRENT APPLICATION NUMBER: PCT/US00/27239

; CURRENT FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 09/411,977

; PRIOR FILING DATE: 1999-10-04

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 2

; LENGTH: 263

; TYPE: PRT

; ORGANISM: Homo sapiens

PCT-US00-27239-2

Query Match 100.0%; Score 1393; DB 1; Length 263;

Best Local Similarity 100.0%; Pred. No. 6.3e-129;

Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAWVQAFVLSNMILAEAYSGGCFWMDNGHLHYREDQTSPPAGLRCLNMDAOGSLASAP 60

DB 1 MLAWVQAFVLSNMILAEAYSGGCFWMDNGHLHYREDQTSPPAGLRCLNMDAOGSLASAP 60

QY 61 VSGAGNHSYCRNPEDPRGRCVSGEAGVPERKPCDELRCPELTSSQALPAFTTEIOEAS 120

Db 61 VSGAGNHSYCRNDEDEPRGWCYVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEAS 120
QY 121 EGPAGAEVQVFAPANALPARSEAAAVQPIVIGISQVRNNSKEKKDLGTLGVIGITMMVT 180
Db 121 EGPAGAEVQVFAPANALPARSEAAAVQPIVIGISQVRNNSKEKKDLGTLGVIGITMMVT 180
QY 181 IIAIGAGIILIGSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIIVDEKTVVHT 240
Db 181 IIAIGAGIILIGSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIIVDEKTVVHT 240
QY 241 SQTVPDQEGSTPLMGQAQTPGA 263
Db 241 SQTVPDQEGSTPLMGQAQTPGA 263

RESULT 2

PCT-US02-05301-164
Sequence 164, Application PC/TUS0205301
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: 70 Human Secreted Proteins
FILE REFERENCE: PS736PCT
CURRENT APPLICATION NUMBER: PCT/US02/05301
CURRENT FILING DATE: 2002-02-21
PRIOR FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: US 60/304,417
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: US 60/270,625
NUMBER OF SEQ ID NOS: 340
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 164
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-05301-164

Query Match 100.0%; Score 1393; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 6.3e-129;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAWQAFVSNMLLAAYVSGGCFWMDNGHLYREDQTSAPGLRCINMLDQSGLASAP 60
Db 1 MLAWQAFVSNMLLAAYVSGGCFWMDNGHLYREDQTSAPGLRCINMLDQSGLASAP 60
QY 61 VSGAGNHSYCRNDEDEPRGWCYVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEAS 120
Db 61 VSGAGNHSYCRNDEDEPRGWCYVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEAS 120
QY 121 EGPAGAEVQVFAPANALPARSEAAAVQPIVIGISQVRNNSKEKKDLGTLGVIGITMMVT 180
Db 121 EGPAGAEVQVFAPANALPARSEAAAVQPIVIGISQVRNNSKEKKDLGTLGVIGITMMVT 180
QY 181 IIAIGAGIILIGSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIIVDEKTVVHT 240
Db 181 IIAIGAGIILIGSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIIVDEKTVVHT 240
QY 241 SQTVPDQEGSTPLMGQAQTPGA 263
Db 241 SQTVPDQEGSTPLMGQAQTPGA 263

RESULT 3

PCT-US98-10728-2
Sequence 2, Application PC/TUS9810728
GENERAL INFORMATION:
APPLICANT: EBERNER, REINHARD
APPLICANT: MOORE, PAUL
APPLICANT: RUBEN, STEVEN
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE

CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US98/10728
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: P379PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US98-10728-2

Query Match 100.0%; Score 1393; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 6.3e-129;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAWQAFVSNMLLAAYVSGGCFWMDNGHLYREDQTSAPGLRCINMLDQSGLASAP 60
Db 1 MLAWQAFVSNMLLAAYVSGGCFWMDNGHLYREDQTSAPGLRCINMLDQSGLASAP 60
QY 61 VSGAGNHSYCRNDEDEPRGWCYVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEAS 120
Db 61 VSGAGNHSYCRNDEDEPRGWCYVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEAS 120
QY 121 EGPAGAEVQVFAPANALPARSEAAAVQPIVIGISQVRNNSKEKKDLGTLGVIGITMMVT 180
Db 121 EGPAGAEVQVFAPANALPARSEAAAVQPIVIGISQVRNNSKEKKDLGTLGVIGITMMVT 180
QY 181 IIAIGAGIILIGSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIIVDEKTVVHT 240
Db 181 IIAIGAGIILIGSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIIVDEKTVVHT 240
QY 241 SQTVPDQEGSTPLMGQAQTPGA 263
Db 241 SQTVPDQEGSTPLMGQAQTPGA 263

RESULT 4

US-09-084-491A-2
Sequence 2, Application US/09084491A
GENERAL INFORMATION:
APPLICANT: MOORE, PAUL A.
APPLICANT: RUBEN, STEVEN M.
APPLICANT: EBERNER, REINHARD
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/084,491A
; FILING DATE: 27-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF378
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; FAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-084-491A-2

Query Match 100.0%; Score 1393; DB 14; Length 263;
Best Local Similarity 100.0%; Pred. No. 6,3e-129;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTAWQAFIVSNMLAEAYSGGCFWMDNGHLYREDQTSAPGRLCLNMLDQSGLASAP 60
DB 1 MLTAWQAFIVSNMLAEAYSGGCFWMDNGHLYREDQTSAPGRLCLNMLDQSGLASAP 60
QY 61 VSGAGNHSYCRNPDEDPGPMCVSGEAGVPEKRPCEDLRCPTTSQALPAFTTEIOEAS 120
DB 61 VSGAGNHSYCRNPDEDPGPMCVSGEAGVPEKRPCEDLRCPTTSQALPAFTTEIOEAS 120
QY 121 EGPADDEVQFAPANALPARSEAAAVQPVIGISQVRNMSKEKDLGTGLGYLGITMMVI 180
DB 121 EGPADDEVQFAPANALPARSEAAAVQPVIGISQVRNMSKEKDLGTGLGYLGITMMVI 180
QY 181 IIAIGAGIILGYSYRGKDLKEQHDQVCEREMQRTPLSAFTNPCEIYDEKTVVHT 240
DB 181 IIAIGAGIILGYSYRGKDLKEQHDQVCEREMQRTPLSAFTNPCEIYDEKTVVHT 240
QY 241 SQTVPDQEGSTPLMGAGTPGA 263
DB 241 SQTVPDQEGSTPLMGAGTPGA 263

RESULT 5
US-09-144-889-2

; Sequence 2 Application US/09144889B
; GENERAL INFORMATION:
; APPLICANT: Earl P. Albone
; APPLICANT: Kristine K. Kikly
; TITLE OF INVENTION: KRINGLE1
; FILE REFERENCE: GH-70249
; CURRENT APPLICATION NUMBER: US/09/144, 889B
; CURRENT FILING DATE: 1998-09-01
; EARLIER APPLICATION NUMBER: 60/056, 032
; EARLIER FILING DATE: 1997-09-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-144-889-2

Query Match 100.0%; Score 1393; DB 15; Length 263;
Best Local Similarity 100.0%; Pred. No. 6,3e-129;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTAWQAFIVSNMLAEAYSGGCFWMDNGHLYREDQTSAPGRLCLNMLDQSGLASAP 60
DB 1 MLTAWQAFIVSNMLAEAYSGGCFWMDNGHLYREDQTSAPGRLCLNMLDQSGLASAP 60

QY 61 VSGAGNHSYCRNPDEDPGPMCVSGEAGVPEKRPCEDLRCPTTSQALPAFTTEIOEAS 120
DB 61 VSGAGNHSYCRNPDEDPGPMCVSGEAGVPEKRPCEDLRCPTTSQALPAFTTEIOEAS 120
QY 121 EGPADDEVQFAPANALPARSEAAAVQPVIGISQVRNMSKEKDLGTGLGYLGITMMVI 180
DB 121 EGPADDEVQFAPANALPARSEAAAVQPVIGISQVRNMSKEKDLGTGLGYLGITMMVI 180
QY 181 IIAIGAGIILGYSYRGKDLKEQHDQVCEREMQRTPLSAFTNPCEIYDEKTVVHT 240
DB 181 IIAIGAGIILGYSYRGKDLKEQHDQVCEREMQRTPLSAFTNPCEIYDEKTVVHT 240
QY 241 SQTVPDQEGSTPLMGAGTPGA 263
DB 241 SQTVPDQEGSTPLMGAGTPGA 263

RESULT 6
US-10-057-951-2

; Sequence 2 Application US/10057951
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-like Protease
; FILE REFERENCE: PF378P1
; CURRENT APPLICATION NUMBER: US/10/057,951
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 09/411,977
; PRIOR FILING DATE: 1999-10-04
; PRIOR APPLICATION NUMBER: US 09/084,491
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: US 60/048,000
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent In Ver. 3.1
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-951-2

Query Match 100.0%; Score 1393; DB 24; Length 263;
Best Local Similarity 100.0%; Pred. No. 6,3e-129;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTAWQAFIVSNMLAEAYSGGCFWMDNGHLYREDQTSAPGRLCLNMLDQSGLASAP 60
DB 1 MLTAWQAFIVSNMLAEAYSGGCFWMDNGHLYREDQTSAPGRLCLNMLDQSGLASAP 60
QY 61 VSGAGNHSYCRNPDEDPGPMCVSGEAGVPEKRPCEDLRCPTTSQALPAFTTEIOEAS 120
DB 61 VSGAGNHSYCRNPDEDPGPMCVSGEAGVPEKRPCEDLRCPTTSQALPAFTTEIOEAS 120
QY 121 EGPADDEVQFAPANALPARSEAAAVQPVIGISQVRNMSKEKDLGTGLGYLGITMMVI 180
DB 121 EGPADDEVQFAPANALPARSEAAAVQPVIGISQVRNMSKEKDLGTGLGYLGITMMVI 180
QY 181 IIAIGAGIILGYSYRGKDLKEQHDQVCEREMQRTPLSAFTNPCEIYDEKTVVHT 240
DB 181 IIAIGAGIILGYSYRGKDLKEQHDQVCEREMQRTPLSAFTNPCEIYDEKTVVHT 240
QY 241 SQTVPDQEGSTPLMGAGTPGA 263
DB 241 SQTVPDQEGSTPLMGAGTPGA 263

RESULT 7
US-10-102-704-2

; Sequence 2 Application US/10102704
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-like Protein
; FILE REFERENCE: PF378C1
; CURRENT APPLICATION NUMBER: US/10/102,704

```
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/084,491
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: 60/048,000
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-102-704-2

Query Match      100.0%; Score 1393; DB 25; Length 263;
Best Local Similarity 100.0%; Pred. No. 6.3e-129;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAWVOAFVLSNMLAEAYSGGCFWMDNGHLYREDQTSPPARGLRCINMLDAOSGLASAP 60
DB 1 MLAWVOAFVLSNMLAEAYSGGCFWMDNGHLYREDQTSPPARGLRCINMLDAOSGLASAP 60
QY 61 VSGAGNHSYCNPDDEPRGPMCVSGEAGVPEKRCEDLRCPETTSQALPAFTTEIOEAS 120
DB 61 VSGAGNHSYCNPDDEPRGPMCVSGEAGVPEKRCEDLRCPETTSQALPAFTTEIOEAS 120
QY 121 ESGGADEVQVFPANALPARSEAAAVQPIGISOVRNMSKEKDLGTLGYVLGITMVI 180
DB 121 ESGGADEVQVFPANALPARSEAAAVQPIGISOVRNMSKEKDLGTLGYVLGITMVI 180
QY 181 IIAIGAGIILGYSYKRGDKLEQHDQKVCEREMORITLPLSAFTNPCEIVDEKTVVHT 240
DB 181 IIAIGAGIILGYSYKRGDKLEQHDQKVCEREMORITLPLSAFTNPCEIVDEKTVVHT 240
QY 241 SOTPVDPQEGSTPLMGOAGTPGA 263
DB 241 SOTPVDPQEGSTPLMGOAGTPGA 263

RESULT 8
US-60-048-000-2
; Sequence 2, Application US/60048000
; GENERAL INFORMATION:
; APPLICANT: EBER, REINHARD
; APPLICANT: MOORE, PAUL
; APPLICANT: RUBEN, STEVE
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/048,000
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKS, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF379PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-60-048-000-2
```

```
Query Match      100.0%; Score 1393; DB 27; Length 263;
Best Local Similarity 100.0%; Pred. No. 6.3e-129;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MLAWVOAFVLSNMLAEAYSGGCFWMDNGHLYREDQTSPPARGLRCINMLDAOSGLASAP 60
DB 1 MLAWVOAFVLSNMLAEAYSGGCFWMDNGHLYREDQTSPPARGLRCINMLDAOSGLASAP 60
QY 61 VSGAGNHSYCNPDDEPRGPMCVSGEAGVPEKRCEDLRCPETTSQALPAFTTEIOEAS 120
DB 61 VSGAGNHSYCNPDDEPRGPMCVSGEAGVPEKRCEDLRCPETTSQALPAFTTEIOEAS 120
QY 121 ESGGADEVQVFPANALPARSEAAAVQPIGISOVRNMSKEKDLGTLGYVLGITMVI 180
DB 121 ESGGADEVQVFPANALPARSEAAAVQPIGISOVRNMSKEKDLGTLGYVLGITMVI 180
QY 181 IIAIGAGIILGYSYKRGDKLEQHDQKVCEREMORITLPLSAFTNPCEIVDEKTVVHT 240
DB 181 IIAIGAGIILGYSYKRGDKLEQHDQKVCEREMORITLPLSAFTNPCEIVDEKTVVHT 240
QY 241 SOTPVDPQEGSTPLMGOAGTPGA 263
DB 241 SOTPVDPQEGSTPLMGOAGTPGA 263
```

```
RESULT 9
US-09-791-537-39562
; Sequence 39562, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB.
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39562
; LENGTH: 263
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-791-537-39562
```

```
Query Match      99.8%; Score 1390; DB 21; Length 263;
Best Local Similarity 99.6%; Pred. No. 1.3e-128;
Matches 262; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MLAWVOAFVLSNMLAEAYSGGCFWMDNGHLYREDQTSPPARGLRCINMLDAOSGLASAP 60
DB 1 MLAWVOAFVLSNMLAEAYSGGCFWMDNGHLYREDQTSPPARGLRCINMLDAOSGLASAP 60
QY 61 VSGAGNHSYCNPDDEPRGPMCVSGEAGVPEKRCEDLRCPETTSQALPAFTTEIOEAS 120
DB 61 VSGAGNHSYCNPDDEPRGPMCVSGEAGVPEKRCEDLRCPETTSQALPAFTTEIOEAS 120
QY 121 ESGGADEVQVFPANALPARSEAAAVQPIGISOVRNMSKEKDLGTLGYVLGITMVI 180
DB 121 ESGGADEVQVFPANALPARSEAAAVQPIGISOVRNMSKEKDLGTLGYVLGITMVI 180
QY 181 IIAIGAGIILGYSYKRGDKLEQHDQKVCEREMORITLPLSAFTNPCEIVDEKTVVHT 240
DB 181 IIAIGAGIILGYSYKRGDKLEQHDQKVCEREMORITLPLSAFTNPCEIVDEKTVVHT 240
QY 241 SOTPVDPQEGSTPLMGOAGTPGA 263
DB 241 SOTPVDPQEGSTPLMGOAGTPGA 263
```

db 241 SOTPVPOEGTTPMLGQAGTPGA 263

```
RESULT 10
US-09-927-796-44
; Sequence 44, Application US/09927796
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scott A.
; APPLICANT: Pan, James
; APPLICANT: Pitt, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Metanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: F2931R1C1
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067411
; PRIOR FILING DATE: 1997-12-03
; PRIOR APPLICATION NUMBER: 60/069862
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/095929
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/097978
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/103396
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/108867
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/112851
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/119965
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: 60/123972
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: 60/133459
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/140650
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/140653
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
```

```
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 60/149395
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/151689
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 08/625328
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 08/710802
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 08/800699
; PRIOR FILING DATE: 1997-02-14
; PRIOR APPLICATION NUMBER: 08/826883
; PRIOR FILING DATE: 1997-03-31
; PRIOR APPLICATION NUMBER: 08/829270
; PRIOR FILING DATE: 1997-03-31
; PRIOR APPLICATION NUMBER: 08/928069
; PRIOR FILING DATE: 1997-09-11
; PRIOR APPLICATION NUMBER: 08/934494
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 09/143068
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: 09/143707
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: 09/151889
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: 09/169104
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: 09/202089
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: 09/254311
; PRIOR FILING DATE: 1999-03-03
; PRIOR APPLICATION NUMBER: 09/304003
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/380138
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/380139
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/403297
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 09/423844
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 09/51133
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 09/511631
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 09/664610
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/665350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/690169
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/690189
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/709238
; PRIOR FILING DATE: 2000-11-18
; PRIOR APPLICATION NUMBER: 09/866034
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 09/872035
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 09/884733
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: 09/886342
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: 09/866028
; PRIOR FILING DATE: 2001-08-25
; PRIOR APPLICATION NUMBER: FCT/US97/05230
; PRIOR FILING DATE: 1997-03-31
; PRIOR APPLICATION NUMBER: FCT/US98/19094
; PRIOR FILING DATE: 1998-09-14
```

Mon Apr 7 07:36:46 2003

us-10-057-951-2.rapm

Page 6

```

1 PRIOR APPLICATION NUMBER: PCT/US98/19330
2 PRIOR FILING DATE: 1998-09-16
3 PRIOR APPLICATION NUMBER: PCT/US98/21407
4 PRIOR FILING DATE: 1998-10-09
5 PRIOR APPLICATION NUMBER: PCT/US98/25108
6 PRIOR FILING DATE: 1998-12-01
7 PRIOR APPLICATION NUMBER: PCT/US99/05028
8 PRIOR FILING DATE: 1999-03-08
9 PRIOR APPLICATION NUMBER: PCT/US99/12252
10 PRIOR FILING DATE: 1999-06-02
11 PRIOR APPLICATION NUMBER: PCT/US99/20111
12 PRIOR FILING DATE: 1999-09-01
13 PRIOR APPLICATION NUMBER: PCT/US99/21090
14 PRIOR FILING DATE: 1999-09-15
15 PRIOR APPLICATION NUMBER: PCT/US99/28313
16 PRIOR FILING DATE: 1999-11-30
17 PRIOR APPLICATION NUMBER: PCT/US99/28301
18 PRIOR FILING DATE: 1999-12-01
19 PRIOR APPLICATION NUMBER: PCT/US99/28634
20 PRIOR FILING DATE: 1999-12-01
21 PRIOR APPLICATION NUMBER: PCT/US99/28565
22 PRIOR FILING DATE: 1999-12-02
23 PRIOR APPLICATION NUMBER: PCT/US00/00219
24 PRIOR FILING DATE: 2000-01-05
25 PRIOR APPLICATION NUMBER: PCT/US00/03565
26 PRIOR FILING DATE: 2000-02-11
27 PRIOR APPLICATION NUMBER: PCT/US00/04341
28 PRIOR FILING DATE: 2000-02-18
29 PRIOR APPLICATION NUMBER: PCT/US00/04342
30 PRIOR FILING DATE: 2000-02-18
31 PRIOR APPLICATION NUMBER: PCT/US00/04414
32 PRIOR FILING DATE: 2000-02-22
33 PRIOR APPLICATION NUMBER: PCT/US00/05841
34 PRIOR FILING DATE: 2000-03-02
35 PRIOR APPLICATION NUMBER: PCT/US00/06884
36 PRIOR FILING DATE: 2000-03-15
37 PRIOR APPLICATION NUMBER: PCT/US00/08439
38 PRIOR FILING DATE: 2000-03-30
39 PRIOR APPLICATION NUMBER: PCT/US00/13705
40 PRIOR FILING DATE: 2000-05-17
41 PRIOR APPLICATION NUMBER: PCT/US00/14941
42 PRIOR FILING DATE: 2000-05-30
43 PRIOR APPLICATION NUMBER: PCT/US00/15264
44 PRIOR FILING DATE: 2000-06-02
45 PRIOR APPLICATION NUMBER: PCT/US00/20710
46 PRIOR FILING DATE: 2000-07-28
47 PRIOR APPLICATION NUMBER: PCT/US00/32678
48 PRIOR FILING DATE: 2000-12-01
49 PRIOR APPLICATION NUMBER: PCT/US01/17800
50 PRIOR FILING DATE: 2001-06-01
51 PRIOR APPLICATION NUMBER: PCT/US01/19652
52 PRIOR FILING DATE: 2001-06-20
53 PRIOR APPLICATION NUMBER: PCT/US01/21066
54 PRIOR FILING DATE: 2001-06-29
55 PRIOR APPLICATION NUMBER: PCT/US01/21735
56 PRIOR FILING DATE: 2001-07-09
57 NUMBER OF SEQ ID NOS: 258
58 SEQ ID NO 44
59 LENGTH: 263

```

| | Query Match | Similarity | Score | DB | Length |
|------------|-------------|--|-------|------------|--------|
| Best Local | 99.8% | 99.6% | 1390 | 23 | 263 |
| Matches | 262 | Conservative | 1 | Mismatches | 0 |
| | | | | Indels | 0 |
| | | | | Gaps | 0 |
| Qy | 1 | MLAAVQAVLVNNMLIAEAYGSGGCGCFEMDNGHLYREDDTSPAPGRLCLNMLDMSGTSAP | 60 | | |
| Db | 1 | MLAAVQAVLVNNMLIAEAYGSGGCGCFEMDNGHLYREDDTSPAPGRLCLNMLDMSGTSAP | 60 | | |
| Qy | 61 | VSAGNHSICNNPDDPRGPMVTCYSGEAGVPEPKRCCEDLRCPETTSQALPATTETIOGAS | 120 | | |
| Db | 61 | VSAGNHSICNNPDDPRGPMVTCYSGEAGVPEPKRCCEDLRCPETTSQALPATTETIOGAS | 120 | | |
| Qy | 121 | EGFGAEEVQVFPANALPASEAAAVQVPIGISQVRNRSKEKKDLGTIGYVLGTTMMVT | 180 | | |

| Db | 121 | EGGGADEVQAPAMNLPARSAAAAPVPGISQVRNNSKKKDJGTGYLGIITMMVI | 180 |
|----|-----|---|-----|
| Qy | 181 | IIAIAAGIIIGYGYKRGKQKHOQKVCEREMORITLPLSAFTNPCEIYDEKVVVHT | 240 |
| Db | 181 | IIAIAAGIIIGYGYKRGKQKHOQKVCEREMORITLPLSAFTNPCEIYDEKVVVHT | 240 |
| Qy | 241 | SGTPVDPQEGSTPLMQAGTPGA | 263 |
| Db | 241 | SGTPVDPQEGSTPLMQAGTPGA | 263 |

```

RESULT 11
US-10-210-951-44
? Sequence 44, Application US/10210951
? GENERAL INFORMATION:
? APPLICANT: Ashkenazi, Avi J.
? APPLICANT: Goddard, Andrew
? APPLICANT: Godowski, Paul J.
? APPLICANT: Gunney, Justin L.
? APPLICANT: Hillan, Kenneth J.
? APPLICANT: Marscek, Scott A.
? APPLICANT: Pan, James
? APPLICANT: Pictl, Robert M.
? APPLICANT: Roy, Margaret Ann
? APPLICANT: Smith, Victoria
? APPLICANT: Stone, Donna M.
? APPLICANT: Watanabe, Colin K.
? APPLICANT: Wood, William I.
? FILE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
? TITLE REFERENCE: P2931R1C1
? CURRENT APPLICATION NUMBER: US/10/210,951
? CURRENT FILING DATE: 2002-08-02
? PRIOR APPLICATION NUMBER: 60/014699
? PRIOR FILING DATE: 1996-04-01
? PRIOR APPLICATION NUMBER: 60/026943
? PRIOR FILING DATE: 1996-09-23
? PRIOR APPLICATION NUMBER: 60/059121
? PRIOR FILING DATE: 1997-07-17
? PRIOR APPLICATION NUMBER: 60/059352
? PRIOR FILING DATE: 1997-09-19
? PRIOR APPLICATION NUMBER: 60/062037
? PRIOR FILING DATE: 1997-10-10
? PRIOR APPLICATION NUMBER: 60/063755
? PRIOR FILING DATE: 1997-10-17
? PRIOR APPLICATION NUMBER: 60/063045
? PRIOR FILING DATE: 1997-10-24
? PRIOR APPLICATION NUMBER: 60/063046
? PRIOR FILING DATE: 1997-10-24
? PRIOR APPLICATION NUMBER: 60/066511
? PRIOR FILING DATE: 1997-11-24
? PRIOR APPLICATION NUMBER: 60/066752
? PRIOR FILING DATE: 1997-11-24
? Remaining Prior Application data removed - See File Wrapper or PALM.
? NUMBER OF SEQ ID NOS: 258
? SEQ ID NO 44
? LENGTH: 263
? TYPE: PRT
? ORGANISM: Homo sapiens
US-10-210-951-44

```

| | Query Match | Similarity | 99.8% | Score 1390 | DB 26 | Length 263 |
|-------------|--------------|--|--------------------|------------|--------|------------|
| Best Local | Similarity | 99.6% | Pred. No. 1,3e-128 | | | |
| Matches 262 | Conservative | 1 | Mismatches 0 | Indels 0 | Gaps 0 | |
| Oy | 1 | MLTAAQAEFLVSNMMLTAAAGSGGCPMNGHLYEDPTSPATACIMWLDNOSGLASP | 60 | | | |
| Db | 1 | MLTAAQAEFLVSNMMLTAAAGSGGCPMNGHLYEDPTSPATACIMWLDNOSGLASP | 60 | | | |
| Oy | 61 | VSGAGNHSICSNPEDPRGPMCVSGEAGVEKPCEDLCRCPTTSGALPAFTTEIOEAS | 120 | | | |
| Db | 61 | VSGAGNHSICSNPEDPRGPMCVSGEAGVEKPCEDLCRCPTTSGALPAFTTEIOEAS | 120 | | | |


```
QY 121 EGGADDEVQFAPANALPARSEAAAQVPVIGISQVRMNSKEKKDLGTLGVLTMMVI 180
| | | | |
DB 121 EGGADDEVQFAPANALPARSEAAAQVPVIGISQVRMNSKEKKDLGTLGVLTMMVI 180
QY 181 ITAIGAGIILGYSYKRGDKLKEQHDQKVCEREMORITLPLSAFTNPCEIVDEKTVVHT 240
| | | | |
DB 181 ITAIGAGIILGYSYKRGDKLKEQHDQKVCEREMORITLPLSAFTNPCEIVDEKTVVHT 240
QY 241 SOTPVDPQEGSTPLMGQAGTPGA 263
| | | | |
DB 241 SOTPVDPQEGSTPLMGQAGTPGA 263

RESULT 12
US-10-211-858-44
; Sequence 44, Application US/10211858
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Guiney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Masters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Pitti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/211,858
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-858-44

Query Match 99.8%; Score 1390; DB 26; Length 263;
Best Local Similarity 99.6%; Pred. No. 1.3e-128;
Matches 262; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 121 EGGADDEVQFAPANALPARSEAAAQVPVIGISQVRMNSKEKKDLGTLGVLTMMVI 180
| | | | |
DB 121 EGGADDEVQFAPANALPARSEAAAQVPVIGISQVRMNSKEKKDLGTLGVLTMMVI 180
QY 181 ITAIGAGIILGYSYKRGDKLKEQHDQKVCEREMORITLPLSAFTNPCEIVDEKTVVHT 240
| | | | |
DB 181 ITAIGAGIILGYSYKRGDKLKEQHDQKVCEREMORITLPLSAFTNPCEIVDEKTVVHT 240
QY 241 SOTPVDPQEGSTPLMGQAGTPGA 263
| | | | |
DB 241 SOTPVDPQEGSTPLMGQAGTPGA 263

RESULT 13
US-10-211-884-44
; Sequence 44, Application US/10211884
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Guiney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Masters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Pitti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/211,884
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-884-44

Query Match 99.8%; Score 1390; DB 26; Length 263;
Best Local Similarity 99.6%; Pred. No. 1.3e-128;
Matches 262; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

Db      61  VSGAGNHSYCRNDEDPGRPCWCVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEAS 120
Qy      121  EGPDADEVQVFAFAPANALPARSEAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMMVI 180
Db      121  EGPDADEVQVFAFAPANALPARSEAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMMVI 180
Qy      181  IIAIGAGIILIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIYDEKTVVHT 240
Db      181  IIAIGAGIILIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIYDEKTVVHT 240
Qy      241  SQTVPDPEGSTPLMGAGTPGA 263
Db      241  SQTVPDPEGSTPLMGAGTPGA 263

```

RESULT 14 US-60-230-435-1580

```

; Sequence 1580, Application US/60230435
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL000768
; CURRENT APPLICATION NUMBER: US/60/230,435
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 2991
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1580
; LENGTH: 263
; TYPE: PR
; ORGANISM: HUMAN
US-60-230-435-1580

```

```

Query Match      99.8%; Score 1390; DB 27; Length 263;
Best Local Similarity 99.6%; Pred. No. 1.3e-128;
Matches 262; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1  MLAMVOAFIVSNMLAEAYSGGCFMNDGHLVREDQTSAPAGLRCLNMLDAGSGLASAP 60
Db      1  MLAMVOAFIVSNMLAEAYSGGCFMNDGHLVREDQTSAPAGLRCLNMLDAGSGLASAP 60
Qy      61  VSGAGNHSYCRNDEDPGRPCWCVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEAS 120
Db      61  VSGAGNHSYCRNDEDPGRPCWCVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEAS 120
Qy      121  EGPDADEVQVFAFAPANALPARSEAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMMVI 180
Db      121  EGPDADEVQVFAFAPANALPARSEAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMMVI 180
Qy      181  IIAIGAGIILIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIYDEKTVVHT 240
Db      181  IIAIGAGIILIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIYDEKTVVHT 240
Qy      241  SQTVPDPEGSTPLMGAGTPGA 263
Db      241  SQTVPDPEGSTPLMGAGTPGA 263

```

```

RESULT 15
US-60-389-987-244
; Sequence 244, Application US/60389987
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Bojin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465P2

```

```

; CURRENT APPLICATION NUMBER: US/60/389,987
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 3025
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 244
; LENGTH: 263
; TYPE: PR
; ORGANISM: Homo sapiens
US-60-389-987-244

```

```

Query Match      99.8%; Score 1390; DB 27; Length 263;
Best Local Similarity 99.6%; Pred. No. 1.3e-128;
Matches 262; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1  MLAMVOAFIVSNMLAEAYSGGCFMNDGHLVREDQTSAPAGLRCLNMLDAGSGLASAP 60
Db      1  MLAMVOAFIVSNMLAEAYSGGCFMNDGHLVREDQTSAPAGLRCLNMLDAGSGLASAP 60
Qy      61  VSGAGNHSYCRNDEDPGRPCWCVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEAS 120
Db      61  VSGAGNHSYCRNDEDPGRPCWCVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEAS 120
Qy      121  EGPDADEVQVFAFAPANALPARSEAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMMVI 180
Db      121  EGPDADEVQVFAFAPANALPARSEAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMMVI 180
Qy      181  IIAIGAGIILIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIYDEKTVVHT 240
Db      181  IIAIGAGIILIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIYDEKTVVHT 240
Qy      241  SQTVPDPEGSTPLMGAGTPGA 263
Db      241  SQTVPDPEGSTPLMGAGTPGA 263

```

```

Search completed: April 7, 2003, 06:57:44
Job time : 195.301 secs

```

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: April 7, 2003, 06:48:47 ; Search time 33.2285 Seconds
(without alignments)
1196.989 Million cell updates/sec

Title: US-10-057-951-2

1393

Sequence: 1 MLAWVQAFVSNMLABAY.....PVDQEGSTPLMGQAGTPGA 263

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 679521 seqs, 151232488 residues

Total number of hits satisfying chosen parameters: 679521

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA New:*
1: /cgn2_6/p/odata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/p/odata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/p/odata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/p/odata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/p/odata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/p/odata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/p/odata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|-----------------------|-------------------|
| 1 | 1384 | 89.4 | 263 | 6 US-10-218-140-6002 | Sequence 6002, Ap |
| 2 | 1120.5 | 80.4 | 264 | 6 US-10-144-779-468 | Sequence 488, Ap |
| 3 | 154.5 | 11.1 | 293 | 7 US-60-452-680-14406 | Sequence 14406, A |
| 4 | 154.5 | 11.1 | 293 | 7 US-60-453-135-8958 | Sequence 8958, Ap |
| 5 | 154.5 | 11.1 | 293 | 7 US-60-453-050-8958 | Sequence 8958, Ap |
| 6 | 154.5 | 11.1 | 516 | 7 US-60-452-680-14405 | Sequence 14405, A |
| 7 | 154.5 | 11.1 | 516 | 7 US-60-453-135-8957 | Sequence 8957, Ap |
| 8 | 154.5 | 11.1 | 527 | 5 US-09-612-314A-51 | Sequence 51, Ap |
| 9 | 154.5 | 11.1 | 527 | 5 PCT-US02-3253-26 | Sequence 26, Ap |
| 10 | 154.5 | 11.1 | 562 | 5 US-09-703-699A-4 | Sequence 4, Ap |
| 11 | 154.5 | 11.1 | 562 | 6 US-10-287-99A-26 | Sequence 26, Ap |
| 12 | 154.5 | 11.1 | 562 | 6 US-60-452-680-14408 | Sequence 14408, A |
| 13 | 154.5 | 11.1 | 562 | 7 US-60-453-135-8960 | Sequence 8960, Ap |
| 14 | 154.5 | 11.1 | 562 | 7 US-60-453-050-8960 | Sequence 8960, Ap |
| 15 | 154.5 | 11.1 | 587 | 5 US-09-949-016-11501 | Sequence 11501, A |
| 16 | 154.5 | 11.1 | 655 | 7 PCT-US02-19017-28 | Sequence 28, Ap |
| 17 | 146.5 | 10.5 | 656 | 7 US-60-453-135-11516 | Sequence 11516, A |
| 18 | 146.5 | 10.5 | 656 | 7 US-60-453-050-11516 | Sequence 11516, A |
| 19 | 146.5 | 10.5 | 656 | 7 US-60-453-135-11516 | Sequence 11516, A |
| 20 | 146 | 10.5 | 322 | 1 PCT-US02-27855-21 | Sequence 21, Ap |
| 21 | 146 | 10.5 | 322 | 1 PCT-US02-27855A-21 | Sequence 21, Ap |
| 22 | 143.5 | 10.3 | 482 | 7 US-60-452-680-14407 | Sequence 14407, A |
| 23 | 143.5 | 10.3 | 482 | 7 US-60-453-135-8959 | Sequence 8959, Ap |
| 24 | 143.5 | 10.3 | 482 | 7 US-60-453-050-8959 | Sequence 8959, Ap |
| 25 | 143.5 | 10.3 | 713 | 5 US-09-949-016-9983 | Sequence 9983, Ap |
| 26 | 143 | 10.3 | 713 | 5 US-09-949-016-9983 | Sequence 9983, Ap |

| | | | | | |
|----|-------|------|-----|-----------------------|-------------------|
| 27 | 143 | 10.3 | 713 | 7 US-60-452-680-22399 | Sequence 22399, A |
| 28 | 143 | 10.3 | 713 | 7 US-60-453-135-13847 | Sequence 13847, A |
| 29 | 143 | 10.3 | 713 | 7 US-60-453-050-13847 | Sequence 13847, A |
| 30 | 142 | 10.2 | 339 | 6 US-10-401-108-3 | Sequence 3, Ap |
| 31 | 142 | 10.2 | 339 | 6 US-10-292-418-11 | Sequence 11, Ap |
| 32 | 142 | 10.2 | 391 | 1 PCT-US02-37879-7 | Sequence 7, Ap |
| 33 | 142 | 10.2 | 391 | 1 PCT-US02-37879-7 | Sequence 7, Ap |
| 34 | 142 | 10.2 | 394 | 6 US-10-304-287-8 | Sequence 8, Ap |
| 35 | 142 | 10.2 | 394 | 6 US-10-304-287-8 | Sequence 8, Ap |
| 36 | 142 | 10.2 | 453 | 5 US-09-291-200A-1 | Sequence 1, Ap |
| 37 | 142 | 10.2 | 458 | 5 US-09-946-893B-4 | Sequence 4, Ap |
| 38 | 142 | 10.2 | 569 | 5 US-09-946-893B-5 | Sequence 5, Ap |
| 39 | 142 | 10.2 | 571 | 5 US-09-946-893B-8 | Sequence 8, Ap |
| 40 | 142 | 10.2 | 576 | 5 US-09-946-893B-6 | Sequence 6, Ap |
| 41 | 142 | 10.2 | 791 | 1 PCT-US02-37879-1 | Sequence 1, Ap |
| 42 | 142 | 10.2 | 791 | 1 US-10-304-287-1 | Sequence 1, Ap |
| 43 | 142 | 10.2 | 810 | 5 US-09-291-200A-5 | Sequence 5, Ap |
| 44 | 142 | 10.2 | 810 | 5 US-09-946-893B-2 | Sequence 2, Ap |
| 45 | 141.5 | 10.2 | 688 | 1 PCT-US02-27855-18 | Sequence 18, Ap |

ALIGNMENTS

```
RESULT 1
US-10-218-140-6002
; Sequence 6002, Application US/10218140
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THEREBY
; FILE REFERENCE: 15966-543 CON
; CURRENT APPLICATION NUMBER: US/10/218,140
; PRIOR FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/540,763
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/127,728
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/127,636
; PRIOR FILING DATE: 1999-04-02
; PRIOR APPLICATION NUMBER: 60/127,607
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 6122
; SOFTWARE: Curator Version 1.0
; SEQ ID NO 6002
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-218-140-6002
Query Match: 39.4%; Score 1384; DB 6; Length 263;
Best Local Similarity: 99.2%; Pred. No. 5.5e-11;
Matches 261; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLAWVQAFVSNMLABAYSGGCFMDNGLHYRDOTSPARGRCIMYDAGSGASAP 60
Db 1 MLAWVQAFVSNMLABAYSGGCFMDNGLHYRDOTSPARGRCIMYDAGSGASAP 60
QY 61 VSAGNHSYCRNPDDPGRPCVYSGAGVPEKRPCEDRCPETTSQALPAFTTEIOEAS 120
Db 61 VSAGNHSYCRNPDDPGRPCVYSGAGVPEKRPCEDRCPETTSQALPAFTTEIOEAS 120
QY 121 EGRPADEVVFPANALPARSEAAVOPVIGISQRYRMSKSKDGLTGVYGIIMVY 180
Db 121 EGRPADEVVFPANALPARSEAAVOPVIGISQRYRMSKSKDGLTGVYGIIMVY 180
QY 181 IIAIGAGIILGYSYRGKDLKEQHDQVCEREMORTTLPISAFNPTCEIIVDEKTVVHT 240
Db 181 IIAIGAGIILGYSYRGKDLKEQHDQVCEREMORTTLPISAFNPTCEIIVDEKTVVHT 240
QY 241 SOTPVDPQEGSTPLMGQAGTPGA 263
Db 241 SOTPVDPQEGSTPLMGQAGTPGA 263
```



```

; APPLICANT: CARGILL, Michele
; APPLICANT: GRUBE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01450
; CURRENT APPLICATION NUMBER: US/60/452,680
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14405
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-452-680-14405

Query Match
Best Local Similarity 11.1%; Score 154.5; DB 7; Length 516;
Best Local Similarity 39.6%; Pred. No. 5.9e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 25 CFMDNGHLVREDQTSPPAGRLCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 75
DB 81 CYEDQGISYRGWTSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGLGNHNYCRNPDR 138
QY 76 DPRGPMCYVSGEAGVPEKRPCEDLRCPEPTS 106
DB 139 DSK-PWCYVF-KAGKYSSEFCSTPACSEGENS 167

RESULT 7
US-60-453-135-8957
; Sequence 8957, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01456
; CURRENT APPLICATION NUMBER: US/60/453,135
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8957
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-8957

Query Match
Best Local Similarity 11.1%; Score 154.5; DB 7; Length 516;
Best Local Similarity 39.6%; Pred. No. 5.9e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 25 CFMDNGHLVREDQTSPPAGRLCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 75
DB 81 CYEDQGISYRGWTSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGLGNHNYCRNPDR 138
QY 76 DPRGPMCYVSGEAGVPEKRPCEDLRCPEPTS 106
DB 139 DSK-PWCYVF-KAGKYSSEFCSTPACSEGENS 167

RESULT 8
US-60-453-050-8957
; Sequence 8957, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01457
; CURRENT APPLICATION NUMBER: US/60/453,050
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
```

```

; SEQ ID NO 8957
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-8957

Query Match
Best Local Similarity 11.1%; Score 154.5; DB 7; Length 516;
Best Local Similarity 39.6%; Pred. No. 5.9e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 25 CFMDNGHLVREDQTSPPAGRLCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 75
DB 81 CYEDQGISYRGWTSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGLGNHNYCRNPDR 138
QY 76 DPRGPMCYVSGEAGVPEKRPCEDLRCPEPTS 106
DB 139 DSK-PWCYVF-KAGKYSSEFCSTPACSEGENS 167

RESULT 9
US-09-612-314A-51
; Sequence 51, Application US/09612314A
; GENERAL INFORMATION:
; APPLICANT: SMITH, RICHARD ANTHONY GODWIN
; APPLICANT: DODD, IAN
; APPLICANT: MOSSAKOWSKA, DANUTA EWA IRENA
; TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
; TITLE OF INVENTION: MEMBRANE-BINDING AGENTS
; FILE REFERENCE: 37945-0004
; CURRENT APPLICATION NUMBER: US/09/612,314A
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 09/214,913
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: PCT/EP97/03715
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: GB 96 148 71.3
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: tissue plasminogen activator
US-09-612-314A-51

Query Match
Best Local Similarity 11.1%; Score 154.5; DB 5; Length 527;
Best Local Similarity 39.6%; Pred. No. 6e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 25 CFMDNGHLVREDQTSPPAGRLCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 75
DB 92 CYEDQGISYRGWTSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGLGNHNYCRNPDR 149
QY 76 DPRGPMCYVSGEAGVPEKRPCEDLRCPEPTS 106
DB 150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGENS 178

RESULT 10
PCT-US02-32263-26
; Sequence 26, Application PC/TUS0232263
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DEFREES, Shawn
; APPLICANT: ZOFF, David
; APPLICANT: BAYER, Robert
; APPLICANT: BOWE, Caryn
; APPLICANT: HAKES, David
; APPLICANT: CHEN, Xi
; TITLE OF INVENTION: REMODELING AND GLYCOCONJUGATION OF PEPTIDES
; FILE REFERENCE: 040853-01-5050WO
```


Job time : 36.2285 secs

```
Db 127 CYEDOGISYRGTWSTASGAECTNM--NSSALAQKPYSGRRPDAIRLGLGNHNYCRNPDR 184
QY 76 DPRGPMCYVSGEAGVPEKRPCEDLRCPEETS 106
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGENS 213
```

RESULT 14

```
US-60-453-135-8960
; Sequence 8960, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CI001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8960
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-8960
```

```
Query Match 11.1%; Score 154.5; DB 7; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.5e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
```

```
QY 25 CFWDNGHLYREDOTSPAPGRCLNWLDAQSGLASAPV-----GAGNHSYCRNPDE 75
Db 127 CYEDOGISYRGTWSTASGAECTNM--NSSALAQKPYSGRRPDAIRLGLGNHNYCRNPDR 184
QY 76 DPRGPMCYVSGEAGVPEKRPCEDLRCPEETS 106
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGENS 213
```

RESULT 15

```
US-60-453-050-8960
; Sequence 8960, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CI001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8960
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-8960
```

```
Query Match 11.1%; Score 154.5; DB 7; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.5e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
```

```
QY 25 CFWDNGHLYREDOTSPAPGRCLNWLDAQSGLASAPV-----GAGNHSYCRNPDE 75
Db 127 CYEDOGISYRGTWSTASGAECTNM--NSSALAQKPYSGRRPDAIRLGLGNHNYCRNPDR 184
QY 76 DPRGPMCYVSGEAGVPEKRPCEDLRCPEETS 106
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGENS 213
```

Search completed: April 7, 2003, 07:03:15

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 06:59:57 ; Search time 15.2003 Seconds

(without alignments)
509.085 Million cell updates/sec

Title: US-10-057-951-2

Perfect score: 1393
Sequence: 1 MLLAWQAFIVSNMLAEVY.....PVDQEGSTPLMGQAGTGA 263

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents AA:*
2: /cgn2_6/prodata/1/aa/5A.COMB.pep:*
3: /cgn2_6/prodata/1/aa/5B.COMB.pep:*
4: /cgn2_6/prodata/1/aa/6A.COMB.pep:*
5: /cgn2_6/prodata/1/aa/6B.COMB.pep:*
6: /cgn2_6/prodata/1/aa/6CTUS.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------------------------|
| 1 | 1393 | 100.0 | 263 | 4 | US-09-411-977-2 Sequence 2, Appl: |
| 2 | 153.5 | 11.7 | 422 | 2 | US-08-811-949-63 Sequence 63, Appl: |
| 3 | 154.5 | 11.1 | 437 | 2 | US-08-811-949-49 Sequence 49, Appl: |
| 4 | 154.5 | 11.1 | 437 | 2 | US-08-811-949-51 Sequence 51, Appl: |
| 5 | 154.5 | 11.1 | 437 | 2 | US-08-811-949-55 Sequence 55, Appl: |
| 6 | 154.5 | 11.1 | 437 | 2 | US-08-811-949-57 Sequence 57, Appl: |
| 7 | 154.5 | 11.1 | 327 | 1 | US-07-609-5108-16 Sequence 16, Appl: |
| 8 | 154.5 | 11.1 | 327 | 2 | US-08-811-949-39 Sequence 39, Appl: |
| 9 | 154.5 | 11.1 | 327 | 2 | PCT-US91-01025A-2 Sequence 2, Appl: |
| 10 | 154.5 | 11.1 | 327 | 6 | 5185259-8 Patent No. 5185259 |
| 11 | 154.5 | 11.1 | 527 | 6 | 5520913-1 Patent No. 5520913 |
| 12 | 154.5 | 11.1 | 546 | 6 | 5200340-6 Patent No. 5200340 |
| 13 | 154.5 | 11.1 | 562 | 2 | US-08-811-949-43 Sequence 43, Appl: |
| 14 | 154.5 | 11.1 | 562 | 2 | US-08-560-098A-50 Sequence 50, Appl: |
| 15 | 154.5 | 11.1 | 562 | 2 | US-08-883-795A-38 Sequence 38, Appl: |
| 16 | 154.5 | 11.1 | 562 | 6 | 5185259-3 Patent No. 5185259 |
| 17 | 154.5 | 11.1 | 562 | 6 | 5200340-2 Patent No. 5200340 |
| 18 | 154.5 | 11.1 | 562 | 6 | 5344773-2 Patent No. 5344773 |
| 19 | 148.5 | 10.7 | 83 | 2 | US-08-811-949-2 Sequence 2, Appl: |
| 20 | 146.5 | 10.5 | 655 | 1 | US-08-148-910-12 Sequence 12, Appl: |
| 21 | 146.5 | 10.5 | 655 | 1 | US-08-448-937A-12 Sequence 12, Appl: |
| 22 | 142 | 10.2 | 160 | 3 | US-08-612-788-35 Sequence 35, Appl: |
| 23 | 142 | 10.2 | 160 | 3 | US-09-066-028-35 Sequence 35, Appl: |
| 24 | 142 | 10.2 | 250 | 3 | US-08-612-788-30 Sequence 30, Appl: |
| 25 | 142 | 10.2 | 250 | 3 | US-09-066-028-30 Sequence 30, Appl: |
| 26 | 142 | 10.2 | 339 | 1 | US-08-248-629A-3 Sequence 3, Appl: |
| 27 | 142 | 10.2 | 339 | 1 | US-08-451-932-3 Sequence 3, Appl: |

| | | | | | |
|----|-----|------|-----|---|-------------------------------------|
| 28 | 142 | 10.2 | 339 | 1 | US-08-452-260-3 Sequence 3, Appl: |
| 29 | 142 | 10.2 | 339 | 2 | US-08-325-785-3 Sequence 3, Appl: |
| 30 | 142 | 10.2 | 339 | 2 | US-08-612-788-3 Sequence 3, Appl: |
| 31 | 142 | 10.2 | 339 | 2 | US-08-608-598B-3 Sequence 3, Appl: |
| 32 | 142 | 10.2 | 339 | 2 | US-08-428-743-3 Sequence 3, Appl: |
| 33 | 142 | 10.2 | 339 | 2 | US-08-886-735-3 Sequence 3, Appl: |
| 34 | 142 | 10.2 | 339 | 3 | US-09-066-028-3 Sequence 3, Appl: |
| 35 | 142 | 10.2 | 339 | 5 | PCT-US95-05107-3 Sequence 3, Appl: |
| 36 | 142 | 10.2 | 352 | 2 | US-08-612-788-40 Sequence 40, Appl: |
| 37 | 142 | 10.2 | 352 | 2 | US-09-066-028-40 Sequence 40, Appl: |
| 38 | 142 | 10.2 | 375 | 4 | US-09-377-250-3 Sequence 3, Appl: |
| 39 | 142 | 10.2 | 375 | 4 | US-09-377-250-2 Sequence 2, Appl: |
| 40 | 142 | 10.2 | 378 | 2 | US-08-612-788-42 Sequence 42, Appl: |
| 41 | 142 | 10.2 | 378 | 3 | US-09-066-028-42 Sequence 42, Appl: |
| 42 | 142 | 10.2 | 378 | 4 | US-09-206-059-1 Sequence 1, Appl: |
| 43 | 142 | 10.2 | 451 | 4 | US-09-377-250-1 Sequence 1, Appl: |
| 44 | 142 | 10.2 | 452 | 4 | US-09-377-250-4 Sequence 4, Appl: |
| 45 | 142 | 10.2 | 790 | 1 | US-08-469-486-54 Sequence 54, Appl: |

ALIGNMENTS

RESULT 1
US-09-411-977-2
; Sequence 2, Application US/09411977
; Patent No. 6372473
; GENERAL INFORMATION:
; APPLICANT: Moore, Paul A.
; APPLICANT: Ruben, Steven M.
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Tissue Plasminogen Activator-like Protease
; FILE REFERENCE: PF378P1
; CURRENT APPLICATION NUMBER: US/09/411,977
; EARLIER PILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: 09/084,491
; EARLIER PILING DATE: 1998-05-27
; EARLIER APPLICATION NUMBER: 60/048,000
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-411-977-2

Query Match 100.0%; Score 1393; DB 4; Length 263;
Best Local Similarity 100.0%; Pred No 1.7e-139;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLAWQAFIVSNMLAEVYSGGCFWDSQSPAPGRLGLMWDQSGTASAP 60
DB 1 MLLAWQAFIVSNMLAEVYSGGCFWDSQSPAPGRLGLMWDQSGTASAP 60
QY 61 VSGAGNYSYRNPDEDPFGWCYSGAGYFEXKPCEDLACPETTSQALPATHEIOAS 120
DB 61 VSGAGNYSYRNPDEDPFGWCYSGAGYFEXKPCEDLACPETTSQALPATHEIOAS 120
QY 121 RSGGADDEVQFAPANALPARESAAYVPIYISQVRMNSYKKDGLTGYLGITMMVY 180
DB 121 RSGGADDEVQFAPANALPARESAAYVPIYISQVRMNSYKKDGLTGYLGITMMVY 180
QY 181 IIAAGAGIIGSYKRGDLKEQHDQVCEREMQITLPLSAFTNPGCIVDEKTVVHT 240
DB 181 IIAAGAGIIGSYKRGDLKEQHDQVCEREMQITLPLSAFTNPGCIVDEKTVVHT 240
QY 241 SQTVPDQEGSTPLMGQAGTGA 263
DB 241 SQTVPDQEGSTPLMGQAGTGA 263

RESULT 2

```

US-08-811-949-63
Sequence 63, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIMA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESS: OBLOH, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLOH, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 472 amino acids
TYPES: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-63
Query Match 11.7% Score 163.5; DB 2; Length 472;
Best Local Similarity 40.0%; Pred No. 7.2e-09;
Matches 38; Conservative 6; Mismatches 38; Indels 13; Gaps 4
Cy 21 GGGCGFNDGHTVREDQNSPARGLRLNLDAGSLASAPVS-----GAGNHSYR 71
Db 33 GARSTEDQGISIRGHTWIASGAECLTN--NSHMLQKFIQGRPRDPIRLGNNHYR 90
Oy 72 NNDDEPRGWCYVSGEAGVPEKPRCEDLACPPTTS 106
Db 91 NEDRDSK-FMCIYF-KAGKYSSEFCSTPACSHGNS 123
RESULT 3
US-08-811-949-49
Sequence 49, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIMA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESS: OBLOH, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLOH, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 472 amino acids
TYPES: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

```

```

1 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
2 CITY: ARLINGTON
3 STATE: VA
4 COUNTRY: USA
5 ZIP: 22202
6 COMPUTER READABLE FORM:
7 MEDIUM TYPE: Floppy disk
8 COMPUTER: IBM PC compatible
9 OPERATING SYSTEM: PC-DOS/MS-DOS
10 SOFTWARE: Patent Release #1.0, Version #1.30
11 CURRENT APPLICATION DATA:
12 APPLICATION NUMBER: US/08/811,949
13 FILING DATE: 05-MAR-1997
14 CLASSIFICATION: 435
15 ATTORNEY/AGENT INFORMATION:
16 NAME: OBLON, NORMAN P.
17 REGISTRATION NUMBER: 24,618
18 REFERENCE/DOCKET NUMBER: 18-966-0
19 TELECOMMUNICATION INFORMATION:
20 TELEPHONE: 703-413-3000
21 TELEFAX: 703-413-2220
22 INFORMATION FOR SEQ ID NO: 49:
23 SEQUENCE CHARACTERISTICS:
24 LENGTH: 437 amino acids
25 TYPE: amino acid
26 TOPOLOGY: linear
27 MOLECULE TYPE: protein
28 US-08-811-949-49
29
30 Query Match 11.1%; Score 154.5; DB 2; Length 437;
31 Best local similarity 39.6%; Pred No 5,8e+08;
32 Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps
33
34 Cy 25 CGMNGHLYRDEOTSPPAGKRCIMLMDGLSAPV-----CGMNYGRNDE 75
35 | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
36 2 CLEOGLSTYRGWTSFAESGAECTNM--NSSALAKPTSSRRPADRIQLGNHYCNDR 59
37 | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
38 Db 76 DRPMQMYGYSGAGGVPEKKRCEDLRCPETTS 106
39 | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
40 60 DSK-PKCYTF-KMGKYSEECFCSTPACSEGS 88
41
42 RESULT 4
43 US-08-811-949-51
44 Sequence 51 Application US/08811949
45 Patent No. 5840531
46 GENERAL INFORMATION:
47 APPLICANT: NIPA, MINBO
48 APPLICANT: SAITO, YOSHIMASA
49 APPLICANT: SASAKI, HITOSHI
50 APPLICANT: HAYASHI, MASAO
51 APPLICANT: NOTANI, JOUJI
52 APPLICANT: KOBAYASHI, MASARAZU
53 TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
54 NUMBER OF SEQUENCES: 67
55 CORRESPONDENCE ADDRESS:
56 ADDRESSEE: OBLON, SEPIYAK, MCCLELLAND, WAIER & NEUSTADT,
57 ADDRESS: P.C.
58 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
59 CITY: ARLINGTON
60 STATE: VA
61 COUNTRY: USA
62 ZIP: 22202
63 COMPUTER READABLE FORM:
64 MEDIUM TYPE: Floppy disk
65 COMPUTER: IBM PC compatible
66 OPERATING SYSTEM: PC-DOS/MS-DOS
67 SOFTWARE: Patent Release #1.0, Version #1.30
68 CURRENT APPLICATION DATA:
69 APPLICATION NUMBER: US/08/811,949
70 FILING DATE: 05-MAR-1997
71 CLASSIFICATION: 435
72 ATTORNEY/AGENT INFORMATION:

```


Patent No. 5326700
GENERAL INFORMATION:
APPLICANT: Berg et al.
TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue PI
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: IN
COUNTRY: U.S.A.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/609,510B
FILING DATE: 19901106
CLASSIFICATION: 435
INFORMATION FOR SEQ. ID NO. 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-609-510B-16

Query Match 11.1%; Score 154.5; DB 1; Length 527;
Best Local Similarity 39.6%; Pred. No. 7.7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 25 CFWDNGHLVREDQTSPPAPGLRCLNWLDAQSLASAPVS-----GAGNHSYCRNPDE 75
DB 92 CYEDGISTRTGWTSTAESGAECTNM--NSSALAKKPYSGRRPDALRLGLGHNHYCRNPDR 149
QY 76 DPRGPMCVYSGEAGVPEKRPCEDLRCPEPTTS 106
DB 150 DSK-PMCYVF-KAGKYSSEFCSTPACSEGNS 178

RESULT 8
US-08-811-949-39
Sequence 39; Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIVA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ. ID NO. 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-39

Query Match 11.1%; Score 154.5; DB 2; Length 527;
Best Local Similarity 39.6%; Pred. No. 7.7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 25 CFWDNGHLVREDQTSPPAPGLRCLNWLDAQSLASAPVS-----GAGNHSYCRNPDE 75
DB 92 CYEDGISTRTGWTSTAESGAECTNM--NSSALAKKPYSGRRPDALRLGLGHNHYCRNPDR 149
QY 76 DPRGPMCVYSGEAGVPEKRPCEDLRCPEPTTS 106
DB 150 DSK-PMCYVF-KAGKYSSEFCSTPACSEGNS 178

RESULT 9
PCT-US91-01025A-2
Sequence 2; Application PC/TUS9101025A
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/01025A
FILING DATE: 19910214
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/486,657
FILING DATE: 1 March 1990
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 454P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/266-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
PCT-US91-01025A-2

Query Match 11.1%; Score 154.5; DB 5; Length 527;
Best Local Similarity 39.6%; Pred. No. 7.7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

OY 25 CFMDGHLHYREQTSPPAGLGLCLMWLDQAQSILAAPVS-----GAGNHSYCRNPDE 75
 : : : : :
 Db 92 CYEDQGYSYRGWSTAESGAECTNW--NSSALAQKPYSGRRPDARIILGLGNHNYCRRPDR 149
 OY 76 DPRGWCYVSGEAGVPEKKRPCEDLRCPEPTTS 106
 : : : : :
 Db 150 DSK-PWCYVF-NAGKYSSSEFCSTPAECSEGS 178

```

RESULT 10
5185259-8
; Patent No. 5185259
; APPLICANT: GOEDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
; VEHAR, GORDON A.
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
; ACTIVATOR
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/489,855
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 12,694
; FILING DATE: 09-FEB-1987
; APPLICATION NUMBER: 483,052
; FILING DATE: 07-APR-1983
; APPLICATION NUMBER: 398,003
; FILING DATE: 14-JUL-1982
; APPLICATION NUMBER: 374,860
; FILING DATE: 05-MAY-1982
; SEQ ID NO.:
; LENGTH: 527
5185259-8

```

```

Query Match      11.1% Score 154.5; DB 6; Length 527;
Best Local Similarity 39.6%; Pred. No. 7.7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4

QY 25 CFWDNGHLLREDDTSPAPGRCLNMLDAOSGLASAPVS-----GAGNHSCENPDE 75
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 92 CYEDQISIRGTWSTLSSGAECTNW--NSSALAKPEYSGRPRDAILRLGLGNHNHYCRNPDR 149

QY 76 DPGFPCVSYSGAAGVEPKRPECEDLRCPETTS 106
   |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 150 DSK-LPWCIVF-KAGKYSSEFCSTPPACSEGNS 178
```

```

RESULT 11
5520913-1
; Patent No. 5520913
; APPLICANT: ANDERSON, STEPHEN BENNETT, WILLIAM F.; BORTSTEIN
; DAVID; HIGGINS, DEBORAH L.; PRONI, NICHOLAS F.; ZOLLER, MARK J.
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR HAVING
; ZYMOGENIC PROPERTIES
; NUMBER OF SEQUENCES: 35
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/88,451
; FILING DATE: 06-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 770,510
; FILING DATE: 03-OCT-1991
; APPLICATION NUMBER: 384,608
; FILING DATE: 24-JUL-1989
; APPLICATION NUMBER: 240,856
; FILING DATE: 02-SEP-1988
; SEQ ID NO:1.
; SEQ LENGTH:527
5520913-1

```

| | | | | |
|-----------------------|-----------------|--------------------|------------|-------------|
| Query Match | 11.1%; | Score 154.5; | DB 6; | Length 527; |
| Best Local Similarity | 39.6%; | Pred. No. 7.7e-08; | | |
| Matches 36; | Conservative 6; | Mismatches 36; | Indels 13; | Gaps 4; |

Oy 25 CFPMNDHKLREDOQTSPAGGLCLMWLDAQGLSAPV-----GAGNHSYRNDE 75
:::
Db 92 CYEDDGISTRGTWSTAESGAECTNW--NSSALQAKFYSGRPPAIRLTGLGNHYCRNDR 148
:::
Oy 76 DPGGPCVYSGEAVPEKRPEDLRCPETTS 106
:::
Db 150 DSK-FMCIYF-KAGKYSEFPCTPACBEGNS 178

```

RESULT 12
5200340-6
: Patent No. 5200340
: APPLICANT: POSTER, DONALD C.;MULVIHILL, EILEEN R.;O'HARA,
: PATRICK J.;PINGEL, KURT;YOSHITAKE, SHINJI
: TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
: ACTIVATORS
: NUMBER OF SEQUENCES: 34
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/53,412
: FILING DATE: 22-MAY-1987
: SEQ ID NO:6:
: LENGTH: 546
5200340-6

```

| | | | | |
|-----------------------|-------|---------------------------------|-----------------------------------|----------------------------------|
| Query Match | 11.1% | Score 154.5 | DB 6 | Length 546 |
| Best Local Similarity | 39.6% | Pred. No. 8.1e-08 | | |
| Matches | 36 | Conservative | 6 | Mismatches 36; Indels 13; Gaps 4 |
| QY | 25 | CPNDGNHLYREDDQTS | PARGLRCLNMLDAGSGLASAPVS----- | GAGNHSRNPDE 75 |
| | | | | |
| Db | 127 | CYEDGGISIRGRWSTAESAGAECTNW-- | NSSLAAQCFYSGRRDPAIRLIGGNHNYGRNPDR | 184 |
| QY | 76 | DPGRGFCVYSGEAGV | PEKRPCEDLACPEPTTS | 106 |
| | | | | |
| Db | 185 | DSK-FWCYVF-YAGKYSSSEFCSTPACSKNS | | 213 |

RESULT 13
 US-08-811-949-43
 Sequence 43, Application US/08811949
 Patent No. 5840533
 GENERAL INFORMATION:
 APPLICANT: NIMA, MINEO
 APPLICANT: SAITO, YOSHIMASA
 APPLICANT: SASAKI, HITOSHI
 APPLICANT: HAYASHI, MASAKO
 APPLICANT: NOTANI, JOUTI
 APPLICANT: KOHAYASHI, MASAKAZU
 TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
 NUMBER OF SEQUENCES: 67
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: P. C.
 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE
 CITY: ARLINGTON
 STATE: VA
 COUNTRY: USA
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/811,949
 FILING DATE: 05-MAR-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: OBLON, NORMAN F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 18-966-0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: April 7, 2003, 08:41:02 ; Search time 19 seconds
(without alignments)
1325.643 Million cell updates/sec

Title: US-10-057-951-2_COPY_2_263
Perfect score: 1388
Sequence: 1 LAMVQAFIVSNMLAEAYG.....PVDPOGSGTPLMGQAGTPGA 262

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR 73:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 154.5 | 11.1 | 562 | 1 UKHUT | t-plasminogen acti |
| 2 | 154 | 11.1 | 559 | 1 A35029 | t-plasminogen acti |
| 3 | 147.5 | 10.6 | 559 | 1 A23941 | t-plasminogen acti |
| 4 | 146.5 | 10.6 | 655 | 1 A46688 | hepatocyte growth |
| 5 | 145.5 | 10.5 | 291 | 2 T38098 | t-plasminogen acti |
| 6 | 142 | 10.2 | 810 | 1 PLHU | plasmin (EC 3.4.21 |
| 7 | 138 | 9.9 | 169 | 2 A40522 | plasmin (EC 3.4.21 |
| 8 | 138 | 9.9 | 433 | 1 UKBAY | u-plasminogen acti |
| 9 | 137 | 9.9 | 716 | 1 JC5061 | macrophage-stimula |
| 10 | 135.5 | 9.8 | 442 | 1 UKPG | u-plasminogen acti |
| 11 | 134 | 9.7 | 431 | 2 JS0599 | t-plasminogen acti |
| 12 | 134 | 9.7 | 477 | 2 JS0598 | t-plasminogen acti |
| 13 | 134 | 9.7 | 477 | 2 JS0598 | t-plasminogen acti |
| 14 | 134 | 9.7 | 477 | 2 JS0598 | t-plasminogen acti |
| 15 | 132 | 9.5 | 394 | 2 JS0600 | macrophage-stimula |
| 16 | 132 | 9.5 | 431 | 1 UKHUT | u-plasminogen acti |
| 17 | 130 | 9.4 | 810 | 2 T46260 | plasmin (EC 3.4.21 |
| 18 | 128.5 | 9.3 | 434 | 1 A35005 | u-plasminogen acti |
| 19 | 128.5 | 9.1 | 433 | 1 JS0500 | u-plasminogen acti |
| 20 | 126.5 | 9.1 | 728 | 1 A60185 | hepatocyte growth |
| 21 | 125.5 | 9.0 | 560 | 1 JC4795 | plasma hyaluronan- |
| 22 | 124.5 | 9.0 | 810 | 2 B30848 | plasmin (EC 3.4.21 |
| 23 | 124 | 8.9 | 622 | 1 TBHU | thrombin (EC 3.4.2 |
| 24 | 124 | 8.9 | 728 | 1 JH0579 | hepatocyte growth |
| 25 | 124 | 8.9 | 812 | 1 PLBO | plasmin (EC 3.4.21 |
| 26 | 123 | 8.9 | 593 | 2 S45281 | coagulation factor |
| 27 | 122.5 | 8.8 | 4548 | 1 S00657 | apoptosis factor |
| 28 | 122 | 8.8 | 603 | 2 S28941 | coagulation factor |
| 29 | 121.5 | 8.8 | 728 | 1 A35644 | hepatocyte growth |

| | | | | | |
|----|-------|-----|------|----------|--------------------|
| 30 | 120 | 8.6 | 433 | 1 UKMS | u-Plasminogen acti |
| 31 | 120 | 8.6 | 790 | 1 PLPG | plasmin (EC 3.4.21 |
| 32 | 119.5 | 8.6 | 477 | 2 JS0597 | t-plasminogen acti |
| 33 | 119 | 8.6 | 710 | 1 I51283 | hepatocyte growth |
| 34 | 118.5 | 8.5 | 615 | 1 KFHU12 | coagulation factor |
| 35 | 118 | 8.5 | 432 | 1 S18932 | u-plasminogen acti |
| 36 | 117 | 8.4 | 625 | 1 TBBO | thrombin (EC 3.4.2 |
| 37 | 116 | 8.4 | 1420 | 2 A32869 | apolipoprotein(a) |
| 38 | 115.5 | 8.3 | 711 | 1 A47136 | macrophage-stimula |
| 39 | 114.5 | 8.2 | 812 | 1 PLMS | plasmin (EC 3.4.21 |
| 40 | 114 | 8.2 | 618 | 2 A35827 | thrombin (EC 3.4.2 |
| 41 | 113 | 8.1 | 558 | 2 UC5878 | plasma hyaluronan- |
| 42 | 112 | 8.1 | 455 | 2 A61545 | plasmin (EC 3.4.21 |
| 43 | 110.5 | 8.0 | 685 | 1 A48289 | neurotrophic recep |
| 44 | 108.5 | 7.8 | 123 | 2 C61545 | plasmin (EC 3.4.21 |
| 45 | 107 | 7.7 | 617 | 2 S10511 | thrombin (EC 3.4.2 |

ALIGNMENTS

RESULT 1
UKHUT
t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human
N/Alternate names: t-PA, tissue plasminogen activator
C/Species: Homo sapiens (man)
C/Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #ext_change 08-Dec-2000
C/Accession: A94004; A23529; J0562; A93293; S02125; A91343; A93951; A91322; A54645; I601
R/NY, T.; Eligh, F.; Lund, B.
Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
A/Title: The structure of the human tissue-type plasminogen activator gene: correlation
A/Reference number: A94004; MUID:84296137; PMID:6089198
A/Accession: A94004
A/Molecule type: DNA
A/Residues: 1-562 <NT>
A/Cross-references: GB:I00141
A/Note: The codon given for residue 93 (ACC) is inconsistent with the authors' translati
R/Fritzer, Degen, S.J.; Rajput, B.; Reich, B.
J. Biol. Chem. 261, 6972-6985, 1986
A/Title: The human tissue plasminogen activator gene.
A/Reference number: A23529; MUID:86196143; PMID:3009482
A/Accession: A23529
A/Molecule type: DNA
A/Residues: 1-562 <DEG>
A/Cross-references: GB:K03021; NID:9339817; PIRN:AA96809.1; PIR:9339818
R/Iitabaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuda, S.; Higashio, K.
Agric. Biol. Chem. 55, 1225-1232, 1991
A/Title: Purification and characterization of tissue plasminogen activator secreted by h
A/Reference number: J0562; MUID:91291340; PMID:1368681
A/Accession: J0562
A/Molecule type: mRNA
A/Residues: 31-562 <ITA>
A/Cross-references: DDBJ:D01096; NID:9220128; PIRN:BA00881.1; PIR:9441174
A/Experimental source: embryonic lung fibroblast IMR-90 cells
A/Note: part of this sequence, including the amino end of the mature protein, was confir
R/Penicic, D.; Holmes, W.E.; Kohr, W.J.; Hartline, R.N.; Vehar, G.A.; Ward, C.A.; Bennett
Nature 301, 214-221, 1983
A/Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche
A/Reference number: A9393; MUID:83115662; PMID:6337343
A/Accession: A93293
A/Molecule type: mRNA
A/Residues: 1-562 <PEN>
A/Cross-references: GB:U00141
A/Experimental source: melanoma cells
R/Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.
Nucleic Acids Res. 16, 5695, 1988
A/Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe
A/Reference number: S02125; MUID:88262579; PMID:3133640
A/Accession: S02125
A/Status: translation not shown
A/Molecule type: mRNA
A/Residues: 1-562 <SAS>
A/Cross-references: EMBL:X07393; NID:937243; PIRN:CAA0302.1; PIR:937244

A;Experimental source: fetal lung cells
R;Agglut. H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Salto, A.; Bando, H.; Okada, K.; Ma
FESS Lett. 189, 145-149, 1985
A;Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen
A;Reference number: A91343; PMID:85285620; PMID:3896853
A;Accession: A91343
A;Molecule type: mRNA
A;Residues: 1-38, 'G', '86-433, 'E', '435-562 <RAG>
A;Experimental source: Detroit 562 cells; ATCC 138
R;Lund, T.; Ny, T.; Randby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
A;Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen acti
A;Reference number: A93951; PMID:83169566; PMID:6572897
A;Accession: A93951
A;Molecule type: mRNA
A;Residues: 251-358 <EDU>
A;Experimental source: melanoma cells
R;Pohl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jorvall, H.
Biochemistry 23, 3701-3707, 1984
A;Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived am
differences.
A;Reference number: A90488; PMID:85000468; PMID:6433976
A;Contents: annotation; melanoma cells, partial sequence of residues 36-562, active and
R;Pohl, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jorvall, H.
FESS Lett. 168, 29-32, 1984
A;Title: Differences between uterine and melanoma forms of tissue plasminogen activator.
A;Reference number: A91322; PMID:84158956; PMID:6538514
A;Accession: A91322
A;Molecule type: protein
A;Residues: 33-45;311-320 <POH>
A;Experimental source: uterus
A;Note in the uterus, cleavage of the activation peptide may also occur after 38-Gln
R;van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.
J. Biol. Chem. 261, 14214-14218, 1986
A;Reference number: A37567; PMID:8703611; PMID:3021732
A;Contents: annotation; fibrin binding site
R;Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Enger
EMBO J. 5, 3525-3530, 1986
A;Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen ac
A;Reference number: A37568; PMID:87161761; PMID:3030730
A;Contents: annotation; fibrin binding site
R;Dodd, I.; Nunn, B.; Robinson, J.H.
Thromb. Haemost. 59, 523-528, 1988
A;Title: Isolation, identification and pharmacokinetic properties of human tissue-type F
A;Reference number: A60902; PMID:8904681; PMID:3142086
A;Contents: annotation; novel forms of expressed recombinant t-PA
R;Harris, T.J.R.; Patel, T.; Marecon, F.A.O.; Little, S.; Emrige, J.S.; Opdenaker, G.;
Mol. Biol. Med. 3, 279-292, 1986
A;Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its exp
A;Reference number: A54645; PMID:86284200; PMID:3030401
A;Accession: A54645
A;Molecule type: mRNA
A;Residues: 1-562 <HAR>
A;Cross-references: GB:M15518; NID:9190031; PIDN:AAA60111.1; PID:9190032
A;Note: Parts of this sequence were confirmed by peptide sequencing
R;Reddy, V.B.; Garrazone, A.J.; Sasak, H.; Wei, C.
DNA 6, 461-472, 1987
A;Title: Expression of human uterine tissue-type plasminogen activator in mouse cells us
A;Accession: 160110; PMID:88054470; PMID:2824147
A;Accession: 160110
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-362 <RES>
A;Cross-references: GB:M18182; NID:9340176; PIDN:AAA36800.1; PID:9340177
R;Fisher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schlenning, W.D.
J. Biol. Chem. 260, 11223-11230, 1985
A;Title: Isolation and characterization of the human tissue-type plasminogen activator B
A;Reference number: I55232; PMID:85289338; PMID:3161893
A;Accession: I55232
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-36 <RE2>
A;Cross-references: GB:M1890; NID:9339837; PIDN:AAA61213.1; PID:9339839

C;Comment: Cleavage by plasmin or trypsin produces two chains held together by a single
C;Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond. It
C;Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat
C;Genetics:
A;Gene: GDB:PLAT
A;Cross-references: GDB:119496; OMIM:173370
A;Map position: 8p12-8p12
A;Intons: 24/3, 39/1, 85/1, 122/1, 180/2, 211/1, 268/2, 297/1, 362/2, 408/1, 455/3, 5
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F;12-32/Domain: signal sequence #status predicted <SIG>
F;24-32/Domain: propeptide #status predicted <PRO>
F;33-552/Product: t-plasminogen activator #status experimental <MAT>
F;33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
F;41-78/Domain: fibronectin type I repeat homology <1F1>
F;86-119/Domain: EGF homology <EGF>
F;127-208/Domain: kringle homology <KR1>
F;215-296/Domain: kringle homology <KR2>
F;311-562/Product: t-plasminogen activator chain B #status experimental <BCH>
F;311-556/Domain: trypsin homology <TRY>
F;41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203,215-296,236-278,267-291,299-
F;152/483/Binding site: carboxydrate (Asn) (covalent) #status experimental
F;219/Binding site: carboxydrate (Asn) (covalent) (partial) #status experimental
F;310-311/Cleavage site: Arg-Ile (plasmin, trypsin) #status experimental
F;357/406/Active site: His, Asp #status predicted
F;513/Active site: Ser #status experimental
Query Match 11.1%; Score 154.5; DB 1; Length 562;
Best Local Similarity 39.6%; Pred. No. 3.4e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
OY 24 CFMDNGLYREQGTSPAGRLCLNLDAGSLASAPVS-----GAGNHSYCRPPE 74
DB 127 CYBQGGSTRGWSFASGAGACTNW--NSSALQKPSGRRPDAIRLGLGNHNCRNPD 184
OY 75 DPGKPCVYSGAGVPEKRPCEDLRCPTTS 105
DB 185 DSK-FMCCYVF-RAGKYSSEFCSTPACSEGNS 213
RESULT 2
t-plasminogen activator (EC 3.4.21.68) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A35029; A31597
R;Feng, P.; Ohlsson, M.; Ny, T.
J. Biol. Chem. 265, 2022-2027, 1990
A;Title: The structure of the rat t-tissue rat tissue-type plasminogen activator gene. Spec
A;Reference number: A35029; PMID:90110448; PMID:2105315
A;Accession: A35029
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-559 <FEN>
A;Cross-references: GB:M11197; NID:9207429; PIDN:AAA42261.1; PID:9207431; GB:J05226
R;Ny, T.; Ohlsson, G.; Haueh, A.J.W.
DNA 7, 671-677, 1988
A;Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activato
A;Reference number: A31597; PMID:89170114; PMID:3148445
A;Accession: A31597
A;Molecule type: mRNA
A;Residues: 1-379, 'K', 381-559 <NYT>
A;Cross-references: GB:M23697; NID:9530159; PIDN:AAA41812.1; PID:9530160
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;11-17/Domain: signal sequence #status predicted <SIG>
F;18-29/Domain: propeptide #status predicted <PRO>
F;30-559/Product: t-plasminogen activator #status predicted <MAT>
F;30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F;38-75/Domain: fibronectin type I repeat homology <1F1>
F;83-116/Domain: EGF homology <EGF>
F;124-205/Domain: kringle homology <KR1>
F;213-294/Domain: kringle homology <KR2>

F:309-559/Domain: t-plasminogen activator chain B #status predicted <BCH>
F:309-553/Domain: trypsin homology <TRY>
F:308-68-66-75, 83-94, 88-105, 107-116, 124-205, 145-187, 176-200, 213-294, 234-276, 265-289, 297-4
F:149, 481/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:308-309/Cleavage site: Arg-116 (plasmin, trypsin) #status predicted
F:355, 404, 510/Active site: His, Asp, Ser #status predicted

Query Match 11.1% Score 154; DB 1; Length 559;
Best Local Similarity 32.4%; Pred. No. 3, 7e-05;

Matches 45; Conservative 14; Mismatches 54; Indels 26; Gaps 6;

24 CFMNGHLYREDQTSPPAPGLRCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 74

124 CFEGQGITTYGWTSTAEKACINW--NSSALSGKPYASAPPAIKLGLGNHNYCRNPDR 181

75 DPGPGMVCYSGAGVEKPCEDLACPE-----TTSQALPAFTTEIOESGPG 123

182 DPK-PWCTVF-KGKTYTTEFCSTPACPKGPTEDCYGKGVTRGTHSTT--SKASCLPW 237

124 ADEVQVAPANALPASEA 142

238 NSMTLIGKTYTARANSOA 256

RESULT 3

t-plasminogen activator (EC 3.4.21.68) precursor - mouse

C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999

A:Accession: A29941; S48205; S48207; S48206

R:Ricklefs, R.J.; Darlow, A.L.; Strickland, S.

J. Biol. Chem. 263, 1563-1569, 1988

A:Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator

A:Reference number: A29941; MUID:86087303; PMID:2826484

A:Accession: A29941

A:Molecule type: mRNA

A:Residues: 1-53 <RIG>

A:Cross-references: GB:U03520; NID:G202109; PIDN:AAA0470.1; PID:G202110

R:LiJinen, H.R.; van Hoel, B.; Beelen, V.; Collen, D.

Eur. J. Biochem. 224, 863-871, 1994

A:Title: Characterization of the murine plasma fibrinolytic system.

A:Reference number: S48202; MUID:95010076; PMID:7523120

A:Accession: S48205

A:Molecule type: protein

A:Residues: 33-37, 'X', 39-40 <LIU>

A:Accession: S48207

A:Molecule type: protein

A:Residues: 309-316 <LIU>

A:Accession: S48206

A:Molecule type: protein

A:Residues: 33-37, 'X', 39-40 <LIU>

C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom

C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-29/Domain: propeptide #status predicted <PRO>

F:30-553/Product: t-plasminogen activator #status predicted <MAT>

F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>

F:38-75/Domain: fibronectin type I repeat homology <IFI>

F:83-116/Domain: EGF homology <EGF>

F:124-205/Domain: kringle homology <KR1>

F:213-294/Domain: kringle homology <KR2>

F:309-559/Product: t-plasminogen activator chain B #status predicted <BCH>

F:308-68-66-75, 83-94, 88-105, 107-116, 124-205, 145-187, 176-200, 213-294, 234-276, 265-289, 297-4

F:149, 481/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:308-309/Cleavage site: Arg-116 (plasmin, trypsin) #status predicted

F:355, 404, 510/Active site: His, Asp, Ser #status predicted

Query Match 10.6% Score 147.5; DB 1; Length 559;
Best Local Similarity 37.0%; Pred. No. 0.00013;
Matches 34; Conservative 11; Mismatches 34; Indels 13; Gaps 4;

24 CFMNGHLYREDQTSPPAPGLRCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 74

DB 124 CFEGQGITTYGWTSTAEKACINW--NSSALSGKPYASAPPAIKLGLGNHNYCRNPDR 181

75 DPGPGMVCYSGAGVEKPCEDLACPE-----TTSQALPAFTTEIOESGPG 123

182 DPK-PWCTVF-KGKTYTTEFCSTPACPKGPTEDCYGKGVTRGTHSTT--SKASCLPW 237

124 ADEVQVAPANALPASEA 142

238 NSMTLIGKTYTARANSOA 256

75 DPGPGMVCYSGAGVEKPCEDLACPE-----TTSQALPAFTTEIOESGPG 123

182 DPK-PWCTVF-KGKTYTTEFCSTPACPKGPTEDCYGKGVTRGTHSTT--SKASCLPW 237

124 ADEVQVAPANALPASEA 142

238 NSMTLIGKTYTARANSOA 256

RESULT 4

hepatocyte growth factor activator (EC 3.4.21.-) precursor [validated] - human

C:Species: Homo sapiens (man)
C:Date: 21-Sep-1993 #sequence_revision 25-Aug-1995 #text_change 08-Dec-2000

A:Accession: A46688

R:Wiyazawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.

J. Biol. Chem. 268, 10024-10028, 1993

A:Title: Molecular cloning and sequence analysis of the cDNA for a human serine protease

A:Reference number: A46688; MUID:93252878; PMID:7683665

A:Accession: A46688

A:Molecule type: mRNA

A:Residues: 1-655 <MTY>

A:Cross-references: DDBJ:U14012; NID:G219680; PIDN:BA01113.1; PID:G219681

A:Experimental source: liver (mRNA); serum (protein)

A:Note: sequence extracted from NCBI backbone (NCBI:131227, NCBI:131228)

A:Note: parts of the sequence, including the amino ends of the heavy and light chains, cc

C:genetics:

A:gene: GDB:HGPAC; HGFA; HGFAV

A:Cross-references: GDB:1954514

A:Map position: 4p16-4p16

C:Function:

A:Description: activates hepatocyte growth factor by specific proteolytic cleavage

A:Superfamily: tissue repair and regeneration

C:Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase

F:1-3/Domain: signal sequence #status predicted <SIG>

F:108-148/Domain: fibronectin type II repeat homology <IF2>

F:164-197/Domain: EGF homology <EG1>

F:202-237/Domain: fibronectin type I repeat homology <IF1>

F:245-278/Domain: EGF homology <EG2>

F:286-367/Domain: kringle homology <KRG>

F:373-407/Product: hepatocyte growth factor activator light chain #status experimental <L

F:408-655/Product: hepatocyte growth factor activator heavy chain #status experimental <H

F:408-641/Domain: trypsin homology <TRY>

F:40, 48, 290, 468, 492, 546/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:164-175, 169-186, 188-197, 228-230, 228-237, 245-256, 250-267, 269-278, 286-367, 307-349, 338-36

F:447, 497, 598/Active site: His, Asp, Ser #status predicted

Query Match 10.6% Score 146.5; DB 1; Length 655;
Best Local Similarity 36.9%; Pred. No. 0.00019;
Matches 41; Conservative 10; Mismatches 41; Indels 19; Gaps 6;

24 CFMNGHLYREDQTSPPAPGLRCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 75

286 CFUGSGTYRKVAVSTASGSLAMNSDLVYOEIHDVS--GGAALALGLDPHAYCKNPUND 344

76 DPGPGMVCYSGAGVEKPCEDLACPE-----TTSQALPAFTTEIOESGPG 121

345 ER-PWCTVF-KGKTYTTEFCSTPACPKGPTEDCYGKGVTRGTHSTT--SKASCLPW 390

138098

RESULT 5

t-plasminogen activator precursor, inactive endothelial splice form - human

C:Species: Homo sapiens (man)

C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999

A:Accession: I38098; S01678

R:Stepper, J.D.; Fong, K.

Nucleic Acids Res. 18, 1086, 1990

A:Title: Variant tissue type plasminogen activator (PLAT) cDNA obtained from human endo

A:Reference number: I38098; MUID:90192128; PMID:1969145

A:Accession: I38098

```

A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-291 <SE>
A:Cross-references: EMBL:X13097; NID:G35282; PIDN:CA31489.1; PID:G35283
C:Comment: For the main splice form, see PIR:UKRUT. This form probably does not have protein
C:Genetics:
A:Gene: GDB:PLAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-8p13
A:Structs: 24/3; 38/1; 85/1; 122/1; 180/2; 211/1; 266/2
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat homology
C:Keywords: alternative splicing; fibroninlys; glycoprotein; kringle
F:1-23/Domain: signal sequence #status predicted <Sig>
F:24-32/Domain: propeptide #status predicted <Pro>
F:33-251/Product: t-plasminogen activator, inactive endothelial splice form #status predicted <FA>
F:41-78/Domain: fibronectin type I repeat homology <EGF>
F:86-119/Domain: EGF homology <EGF>
F:121-208/Domain: kringle homology <KR1>
F:215-291/Domain: kringle homology #status atypical <KR2>
F:41-73; 69-78; 86-97; 91-108; 110-119; 127-208; 148-190; 179-203/Diulfide bonds: #status predicted
Query Match 10.5%; Score 145.5; DB 2; Length 291;
Best Local Similarity 38.5%; Pred. No. 9e-05;
Matches 35; Conservative 6; Mismatches 37; Indels 13; Gaps 4;
OY 24 CFFNDNHLVREDOGTSPAPGLRCILNMLDASGLASPVG-----GAGNNSYRRNPD 74
DB 127 CYEDQGISIRGWSTVNSGAECTVW--NSALDAQNVAISGRAPPAIRLISGNHNYGRNPPDR 184
OY 75 DPRGFWCYISGEAGVFEKRRPCEDLCRPETTS 105
DB 185 DSK-PWCYVF-RAGKYSSSEFCSTPACSENS 213

```

RESULT 6

plasma [EC 3.4.21.7] precursor [validated] - human
N:Alternate names: plasminogen precursor [misnomer]
N:Contains: angiotensin, microplasmin, plasminogen
C:Species: Homo sapiens (man)
C:Cdate: 24-Apr-1986 #sequence, revision 02-Dec-1994 #ext change 15-Sep-2000
C:Accession: A135229; E52242; A6646; I6278; I64609; S0735; A00929; A04627; A04625; A04626
R:Reisen, F.E.; Matzen, M.R.; Ichimose, A.; Davie, E.W.
J. Biol. Chem. 265, 6104-6111, 1990
A>Title: Characterization of the gene for human plasminogen, a key proenzyme in the fibrinolysis system
A:Reference number: A135229; PMID:50202879; PMID:2318648
A:Accession: A135229
A:Molecule type: DNA
A:Residues: 1-810 <PER>
A:Cross-references: GB:J05286; GB:M34276; NID:G190064; PIDD:AAA60113.1; PID:g387026
A:Experimental source: leukocyte, lung fibroblast
R:Margheriti, V., Bruno, L., Pontoglio, M.; Cadiani, G.; Meroni, G.; Ottolenghi, S.; Tosi, R.
Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990
A>Title: Definition of the transcription initiation site of human plasminogen gene in 1 kb genomic DNA
A:Reference number: J52242; PMID:91097523; PMID:12269308
A:Accession: J52242
A>Status: translated from GB/EMBL/DDBC
A:Molecule type: DNA
A:Residues: 1-16 <MAL>
A:Cross-references: GB:M62890; NID:G190092; PIDD:AA36454.1; PID:g553613
R:Forgeren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heiden, L.O.
FEBS Lett. 213, 254-260, 1987
A>Title: Molecular cloning and characterization of a full-length cDNA clone for human plasminogen activator
A:Reference number: A26646; PMID:87162490; PMID:3030813
A:Accession: A26646
A:Molecule type: mRNA
A:Residues: 1-471,'D',473-810 <FOR>
A:Cross-references: GB:X05199; NID:g35530; PIDD:CAA26831.1; PID:g35531
A:Experimental source: liver
R:Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A>Title: Characterization of a complementary deoxyribonucleic acid coding for human and bovine tissue-type plasminogen activator
A:Reference number:145961; PMID:85023311; PMID:6148961

A:Accession: I62738
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 292-471, 'D', 473-810 <MAL2>
A:Cross-references: GB:K02922; NID:G190112; P1DN:AAA60124.1; P1D:G387031
A:Accession: I64609
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 367-419 <MAL3>
A:Cross-references: GB:K02923; NID:G190110; P1DN:AAA60123.1; P1D:G190111
P:Brunsholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, B.E.; Leger, W.; Manneberg, M.,
Eur. J. Biochem. 114, 445-470, 1981
A:Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,
A:Reference number: S03735; MUID:8121097; PMID:7238497
A:Molecule type: protein
A:Residues: 20-71, 'E', 73-85 <BRU>
R:Scrup-Jensen, L.; Petersen, T.B.; Magnusson, S.
submitted to the Atlas, July 1977
A:Reference number: A00929
A:Accession: A00929
A:Molecule type: protein
A:Residues: 20-71, 'E', 73-85, 87-106, 'D', 108-360, 'E', 362-810 <SOT>
R:Wiman, B.
Eur. J. Biochem. 76, 129-137, 1977
A:Title: Primary structure of the B-chain of human plasmin.
A:Reference number: A04627; MUID:7722545; PMID:142009
A:Accession: A04627
A:Molecule type: protein
A:Residues: 581-810 <W1>
R:Wiman, B.; Wallen, P.
Eur. J. Biochem. 50, 489-494, 1975
A:Title: Structural relationship between "glutamic acid" and "lysine" forms of human pla
A:Reference number: A04625; MUID:75093329; PMID:1229332
A:Accession: A04625
A:Molecule type: protein
A:Residues: 20-50 'Q', 51-71, 'B', 73-85, 87-100 <W2>
R:Wiman, B.; Wallen, P.
Eur. J. Biochem. 58, 539-547, 1975
A:Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen chate
A:Reference number: A04626; MUID:76045692; PMID:126863
A:Accession: A04626
A:Molecule type: protein
A:Residues: 483-507, 'E', 509-604 <W3>
R:Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.
Eur. J. Biochem. 249, 1631-1633, 1973
A:Title: The primary structure of human plasminogen. II. The histidine loop of human pl
A:Reference number: A92125; MUID:73149248; PMID:4694729
A:Contents: annotation; active site
R:Groskopf, W.R.; Summaria, L.; Robbins, K.C.
Eur. J. Biochem. 244, 3599-3597, 1969
A:Title: Studies on the active center of human plasmin. Partial amino acid sequence of a
A:Reference number: A92048; MUID:69234739; PMID:4240117
A:Contents: annotation; active site
R:Trexler, M.; Vail, Z.; Patchy, L.
Eur. J. Biochem. 257, 7401-7406, 1982
A:Title: Structure of the omega-amino-carboxylic acid-binding sites of human plasminogen.
A:Reference number: A92382; MUID:82213905; PMID:6919539
A:Contents: annotation; omega-amino-carboxylic acid binding sites
R:Vail, Z.; Patchy, L.
Eur. J. Biochem. 259, 13690-13694, 1984
A:Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential
A:Reference number: A92459; MUID:85054794; PMID:6096526
R:Go, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Sehnemel, S.; McCance, S.G.;
J. Biol. Chem. 271, 29461-29467, 1996
A:Title: Kissing domains of human angiotensin. Characterization of the anti-proliferative
A:Reference number: A58611; MUID:97067211; PMID:8910613
A:Contents: annotation
R:Ujnen, H.R.; Ugwu, F.; Bini, A.; Collen, D.
Biochemistry 31, 4699-4702, 1998
A:Title: Generation of an angiotensin-like fragment from plasminogen by streptolysin-1 (MS

C:Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999
A:Accession: S14687, S08651
R:Au, Y.P.T., Wang, R.W., Clower, A.W.
N:Acid: Nucleotide and deduced amino acid sequence of baboon urokinase-type plasminogen
A:Title: Nucleotide and deduced amino acid sequence of baboon urokinase-type plasminogen
A:Reference number: S14687; MUID:90287734; PMID:2113276
A:Accession: S14687
A:Molecule type: mRNA
A:Residues: 1-433 <ADV>
C:Cross-references: EMBL:X51935; NID:G938130; PIDN:CAA36200.1; PID:G38131
C:Keywords: urokinase-type plasminogen activator; BGF homology; kringle homology; try
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-120/Domain: signal sequence #status predicted <SIG>
F:121-176/Product: plasminogen activator chain A #status predicted <ACH>
F:30-61/Domain: BGF homology <BGF>
F:169-150/Domain: kringle homology <KR3>
F:178-433/Product: plasminogen activator chain B #status predicted <BCH>
F:178-431/Domain: trypsin homology <TRY>
F:167-288, 208-224, 216-287, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted
F:1223, 274, 378/Active site: His, Asp, Ser #status predicted
F:1223/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.9%; Score 138; DB 1; Length 433;
Best Local Similarity 32.7%; Pred No. 0.00061; Mismatches 16; Gaps 4;
Matches 32; Conservative 14; Mismatches 36; Indels 16; Gaps 4;

OY 24 CFMDNGHLYREDQTSPPAPGLRCLINWDA-----QSGLASAPVSGAGNHSYCRNPDEDP 76
Db 69 CYBENGHFRYKRAKSTDTMGKSCLAANWATVLTQYHAHSDMLQLGLGKHNYCRNP-NR 127
OY 77 RGPWCYVSGEAGVPEK-----RPCEDLRCPEETSOAL 108
Db 128 RRPWCYV--QVGLKQRYQECMVHNCADGKKKSPPEHL 163

RESULT 9
JC5061
macrophage-stimulating protein 1 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 16-Jun-2000
A:Accession: JC5061
R:Ohnishi, K., Iwama, A., Matsuno, K., Ezaki, T., Sakamoto, O., Hamauchi, I., Takasu, I.
Biochem. Biophys. Res. Commun. 227, 273-280, 1996
A:Title: Molecular cloning of rat macrophage-stimulating protein and its involvement in
A:Reference number: JC5061; MUID:97011126; PMID:885136
A:Accession: JC5061
A:Molecule type: mRNA
A:Residues: 1-716 <ONS>
C:Cross-references: EMBL:X55096; NID:g1669718; PIDN:CAA64473.1; PID:g1669719
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C:Keywords: hepatocyte growth factor; kringle homology; trypsin homology
C:Keywords: duplication; glycoprotein; growth factor; kringle
F:1-11/Domain: signal sequence #status predicted <SIG>
F:132-488, 489-716/Product: macrophage-stimulating protein 1 #status predicted <MNT>
F:132-488/Domain: macrophage-stimulating protein 1 alpha chain #status predicted <ACH>
F:118-186/Domain: kringle homology <KR11>
F:191-268/Domain: kringle homology <KR12>
F:1293-370/Domain: kringle homology <KR13>
F:1378-457/Domain: kringle homology <KR14>
F:1488-716/Domain: macrophage-stimulating protein 1 beta chain #status predicted <BCH>
F:1489-709/Domain: trypsin homology <TRY>
F:172,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.9%; Score 137; DB 1; Length 716;
Best Local Similarity 27.3%; Pred. No. 0.0013; Mismatches 56; Indels 60; Gaps 8;
Matches 48; Conservative 12; Mismatches 56; Indels 60; Gaps 8;

OY 1 GAAVQAFVSNMMLAAYSG-----GCFMDNGHLYREDQTSPPAPGLRCLINW 48
Db 80 LLEPWTQ-----HSRAQLHHSSLCIDLPFKQKDVYRTITMNGHSYRGVYAKRTADGLPCQAW 134
OY 49 ---LDAQGLASAPVSGAGNHSYCRNPDEDPGRCWCYVS----- 84

Db 135 SRPRNPXHTPTPRXKGL--EENPCRNPDGDP--PWPWCCTTNRSRFQSCGAKSCRAVCTW 193

QY 85 -----GEAGVPEK-RPCE--DLRCPET-----TSALPAFTTEQAESECP 122

Db 194 CNGEDYRGEVDVLTSGRECCQMDLQHPHSHFPEKFPDKALQNYCRNPDASERP 249

RESULT 10

UKPG

U-plasminogen activator (EC 3.4.21.73) precursor - pig

N.Alternate names: uPA

C.Species: Sus scrofa domestica (domestic pig)

C.Date: 04-Dec-1996 #sequence_revision 17-Mar-1997 #ext_change 07-Aug-1998

C.Accession: A00932

R.Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.

Nucleic Acids Res. 15, 9525-9541, 1984

A.Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.

A.Reference number: A00932; MUID:85087954; PMID:6096832

A.Accession: A00932

A.Molecule type: DNA

A.Residues: 1-240, 'H', 242-442 <NAG1>

A.Experimental source: kidney cell line LLC-PK1

R.Nagamine, Y.

submitted to the Protein Sequence Database, December 1986

A.Reference number: A37566

A.Contents: annotation; correction to residue 241

C.Genetics:

A.Introns: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3

C.Keywords: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C.Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase

F.1-20/Domain: signal sequence #status predicted <Sig>

F.21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>

F.233-64/Domain: EGF homology <EGF>

F.72-153/Domain: kringle homology <KRG>

F.190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BGH>

F.119-430/Domain: trypsin homology <TRY>

F.152/Binding site: carbonylate (asn) (covalent) #status predicted

F.117-310, 220-236, 228-239, 324-393, 358-372, 383-411/Disulfide bonds: #status predicted

F.233, 286, 387/Active site: His, Asp, Ser #status predicted

Query Match 9.8%; Score 135.5; DB 1; Length 442;

Best Local Similarity 36.9%; Pred. No. 0.001;

Matches 38; Conservative 7; Mismatches 37; Indels 21; Gaps 5;

QY 24 CFWDGHHYREDQSPAPGLKCLNWLDAQSL---ASAPVS---GAGNNSYCRNPDEP 76

Db 72 CFEQNGSHYRKANTNNGRCRLPMNSATVLLNTYHAHRDALQGLGKKNYCRND-NQ 130

QY 77 RGPWCYVS-----GEAGVP-----EKRPCEDLRCPEFTSQ 106

Db 131 RRPWCYVQVGLKQLVQECWPNCSGSGSHRPAYDGKNPSTPE 173

RESULT 11

JS0599

E-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat

N.Alternate names: tissue plasminogen activator

C.Species: Desmodus rotundus (common vampire bat)

C.Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #ext_change 16-Jul-1999

C.Accession: JS0599

R.Kreitschmar, J.; Haendler, B.; Langer, G.; Boifol, W.; Bringmann, P.; Alagon, A.; Donn

Gene 105, 229-337, 1991

A.Title: The plasminogen activator family from the salivary gland of the vampire bat Des

A.Reference number: JS0597; MUID:92039036; PMID:1937019

A.Accession: JS0599

A.Molecule type: mRNA

A.Residues: 1-431 <KRA>

A.Cross-references: GB:63989; NID:g166076; PIDN:AAA1594.1; PID:g166077

C.Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase

F.1-21/Domain: signal sequence #status predicted <Sig>

F.22-36/Domain: propeptide #status predicted <PRO>


```

Db      80 LLPTQ-----HSLHTQLYHSLCHLFOKQDYVHTCTIMDNGSVYRGTVARTAGILPCQAW 134
QY      49 ---LDAOSGLASAPVSGAGNHSYCRNPDEDPGPMCYVS-----84
Db      135 SRPFNDHKYTPPTPKGL--EENFCRNPDGDRGPMCYTNRSVRFQSCGIKTCREAVCVL 193
QY      85 -----GEAGVPEK-RPCE--DLRCPEPT 103
Db      194 CNGEDYRGVEVDVTESGRECCQWDLQPHS 222

```

RESULT 15

```

JS0600
t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
N/Alternate names: tissue plasminogen activator
C/Species: Desmodus rotundus (common vampire bat)
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C/Accession: JS0600
R/Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Dot
Gene 105, 229-237, 1991
A/TITLE: The plasminogen activator family from the salivary gland of the vampire bat Des
A/Reference number: JS0597; MUID:92039036; PMID:1937019
A/Accession: JS0600
A/Molecule type: mRNA
A/Residues: 1-194 <KRA>
A/Cross-references: GB:M63990; NID:G166078; PIDN:AAA1595.1; PID:G166079
A/Note: The authors translated the codon ATC for residue 75 as Thr
C/Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F/1-21/Domain: signal sequence #status predicted <SIG>
F/12-36/Domain: propeptide #status predicted <PRO>
F/37-394/Product: plasminogen activator gamma #status predicted <PLA>
F/45-126/Domain: kringle homology <KR>
F/143-388/Domain: trypsin homology <TRY>
F/445-126/Domain: trypsin homology <TRY>
F/445-126/Domain: trypsin homology <TRY>
F/142-143/Cleavage site: His-Ser (plasmin) #status predicted
F/189-238/Active site: His, Asp, Ser #status predicted
F/315/Binding site: carbohydrate (Aen) (covalent) #status predicted

```

Query Match

9.5%; Score 132; DB 2; Length 394;

Best Local Similarity 32.7%; Pred. No. 0.0017;

Matches 33; Conservative 9; Mismatches 39; Indels 20; Gaps 5;

```

QY      16 AEAYS--GGCFWNGHLIYREDQTSAPGLRCLMW-----LDAOSGLASAPVSGAGNH 66
Db      35 SRAYGDPRATCYKQGVYRGVTWSTESGACQICIMNNSNLIIRTYNGRMPEAVKGLGNH 94
QY      67 SYCRNPDEDPGPMCYV-----SGEAGVPEKRPCEDLRC 100
Db      95 NYCRRNPDGASK-PWCYIKARKFTSESQV---VCSKATC 131

```

Search completed: April 7, 2003, 08:43:53
Job time : 21 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 08:34:32 ; Search time 11 Seconds
(without alignments)
987.891 Million cell updates/sec

Title: US-10-057-951-2_COPY_2_263

Perfect score: 1388
Sequence: 1 LLAWQAFIVSNMLAEAVG.....PYDPOEGSTPLMGQAGTPCA 262

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|---------------------|
| 1 | 160.5 | 11.6 | 566 | 1 TPA_BOVIN | Q28198 bos taurus |
| 2 | 154.5 | 11.1 | 562 | 1 TPA_HUMAN | P00750 homo sapien |
| 3 | 154 | 11.1 | 559 | 1 TPA_RAT | P19637 rattus norv |
| 4 | 150 | 10.8 | 653 | 1 HGFA_MOUSE | Q97098 mus musculu |
| 5 | 147.5 | 10.6 | 559 | 1 TPA_MOUSE | P11214 mus musculu |
| 6 | 146.5 | 10.6 | 655 | 1 HGFA_HUMAN | O04756 homo sapien |
| 7 | 142 | 10.2 | 810 | 1 PLMN_HUMAN | P00747 homo sapien |
| 8 | 138 | 9.9 | 169 | 1 PLMN_RAT | O01177 rattus norv |
| 9 | 138 | 9.9 | 433 | 1 UROK_PAPCY | P16227 papio cynoc |
| 10 | 135.5 | 9.8 | 442 | 1 UROK_PIG | P04185 sus scrofa |
| 11 | 134 | 9.7 | 431 | 1 URTB_DESRO | P98121 desmodus ro |
| 12 | 134 | 9.7 | 477 | 1 URT2_DESRO | P15638 desmodus ro |
| 13 | 134 | 9.7 | 716 | 1 HGFL_MOUSE | P26928 mus musculu |
| 14 | 132 | 9.5 | 394 | 1 URTG_DESRO | P49150 desmodus ro |
| 15 | 132 | 9.5 | 431 | 1 UROK_HUMAN | P00749 homo sapien |
| 16 | 130 | 9.4 | 810 | 1 PLMN_ERIEU | O29485 erinaceus e |
| 17 | 128.5 | 9.3 | 434 | 1 UROK_CHICK | P15120 gallus galli |
| 18 | 127 | 9.1 | 433 | 1 UROK_BOVIN | O05589 bos taurus |
| 19 | 126.5 | 9.1 | 728 | 1 HGF_MOUSE | O08048 mus musculu |
| 20 | 124.5 | 9.0 | 810 | 1 PLMN_MACMU | P12545 macaca mula |
| 21 | 124 | 8.9 | 622 | 1 THRB_HUMAN | P00734 homo sapien |
| 22 | 124 | 8.9 | 728 | 1 HGF_HUMAN | P11210 homo sapien |
| 23 | 124 | 8.9 | 812 | 1 PLMN_BOVIN | P06868 bos taurus |
| 24 | 123 | 8.9 | 593 | 1 FAI2_BOVIN | P98140 bos taurus |
| 25 | 122.5 | 8.8 | 458 | 1 APOA_HUMAN | P08519 homo sapien |
| 26 | 122 | 8.8 | 603 | 1 FAI2_CAVPO | O04962 cavia porce |
| 27 | 121.5 | 8.8 | 728 | 1 HGF_RAT | P17945 rattus norv |
| 28 | 120.5 | 8.7 | 333 | 1 PLMN_CANFA | P80009 canis famul |
| 29 | 120 | 8.6 | 433 | 1 UROK_MOUSE | P06869 mus musculu |
| 30 | 120 | 8.6 | 790 | 1 PLMN_PIG | P06867 sus scrofa |
| 31 | 119.5 | 8.6 | 477 | 1 URT1_DESRO | P98119 desmodus ro |
| 32 | 118.5 | 8.5 | 615 | 1 FAI2_HUMAN | P00748 homo sapien |
| 33 | 118 | 8.5 | 432 | 1 UROK_RAT | P29598 rattus norv |

| | | | | | |
|----|-------|-----|------|--------------|--------------------|
| 34 | 117 | 8.4 | 625 | 1 THRB_BOVIN | P00735 bos taurus |
| 35 | 116 | 8.4 | 1420 | 1 APOA_MACMU | P14417 macaca mula |
| 36 | 115.5 | 8.3 | 711 | 1 HGFL_HUMAN | P26927 homo sapien |
| 37 | 114.5 | 8.2 | 812 | 1 PLMN_MOUSE | P20918 mus musculu |
| 38 | 114 | 8.2 | 473 | 1 KREM_MOUSE | Q99443 mus musculu |
| 39 | 114 | 8.2 | 473 | 1 KREM_RAT | Q92484 rattus norv |
| 40 | 114 | 8.2 | 475 | 1 KREM_HUMAN | Q96m48 homo sapien |
| 41 | 114 | 8.2 | 618 | 1 THRB_MOUSE | P19221 mus musculu |
| 42 | 109.5 | 7.9 | 1709 | 1 SN_HUMAN | Q9bzz2 homo sapien |
| 43 | 107 | 7.7 | 617 | 1 THRB_RAT | P18292 rattus norv |
| 44 | 103.5 | 7.5 | 343 | 1 PLMN_SHEEP | P1286 ovis aries |
| 45 | 101.5 | 7.3 | 325 | 1 PLMN_PETWA | P33574 petromyzon |

ALIGNMENTS

```

RESULT 1
ID TPA_BOVIN STANDARD; PRT; 566 AA.
AC Q28198;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
RT and tPA.";
RU Int. Dairy J. 5:605-617(1995).
CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMAGEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -!- PTM: PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XIa.
CC -!- MISCELLANEOUS: BINDS TO THE KININASE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X85800; CA59795.1; --
CC HSP: P00750; 1RTR.
CC MEROPS: S01.232; --
CC InterPro: IPR001314; Chymotrypsin.
CC InterPro: IPR000561; BGF-like.
CC InterPro: IPR000083; Fibronctn.

```



```

DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR Prodom; PD000395; Kringle; 2.
DR SMART; SM00184; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_spec; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00700; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
KW Plasma; kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 21
FT PROPEP 22 33
FT CHAIN 34 566
FT CHAIN 34 314
FT CHAIN 315 566
FT DOMAIN 40 82
FT DOMAIN 83 121
FT DOMAIN 128 209
FT DOMAIN 219 300
FT DOMAIN 315 566
FT ACT_SITE 361 361
FT ACT_SITE 410 410
FT ACT_SITE 517 517
FT DISULFID 42 72
FT DISULFID 70 79
FT DISULFID 87 98
FT DISULFID 92 109
FT DISULFID 111 120
FT DISULFID 128 209
FT DISULFID 149 191
FT DISULFID 180 204
FT DISULFID 219 300
FT DISULFID 240 282
FT DISULFID 271 295
FT DISULFID 303 434
FT DISULFID 346 362
FT DISULFID 354 423
FT DISULFID 448 523
FT DISULFID 480 496
FT DISULFID 513 541
FT CARBOHYD 153 153
FT CARBOHYD 487 487
SQ SEQUENCE 566 AA; 63701 MW; 2EB6B84E3276C3 CRC64;

Query Match 11.6%; Score 160.5; DB 1; Length 566;
Best Local Similarity 35.5%; Pred. No. 3.4e-06;
Matches 38; Conservative 11; Mismatches 45; Indels 13; Gaps 4;
QY 15 LAEAYSGGCGFWMGNHLYREDQTSPPAGRLCLNWLDAQSGIAPVY-----GAGN 65
DB 119 LCEIDATATATCYKQGVAVRGWSTAESGARCAMW--NSSSLAKPYSGRRPNAIRLGLGN 176
QY 66 HSYCRNDEDEPRGPHWCYSGEAGYPERKRPEDIRCPETTSQALPAFT 112
DB 177 HNYCRNDDQSK-FWCVF-KAGRTISEFCSTPAKAKVAEEDDCYT 221

```

```

TPA_HUMAN
ID TPA_HUMAN STANDARD; PRT; 562 AA.
AC P00750; Q15103;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (TPA)
DE (t-PA) (t-plasminogen activator) (Ateplase) (Retelase).
GN PLAT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=83115262; PubMed=6337343;
RA Penica D., Holmes W.E., Kohr W.J., Harkins R.N., Vehar G.A., H.L.,
RA Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,
RA Coedel D.V., Collen D.;
RT "Cloning and expression of human tissue-type plasminogen activator
RT cDNA in E. coli.";
RL Nature 301:214-221 (1983).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Retal lung;
RX MEDLINE=86262579; PubMed=3133640;
RA Sasaki H., Saito Y., Hayaashi M., Otsuka K., Niwa M.;
RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA
RT from human fetal lung cells.";
RL Nucleic Acids Res. 16:5695-5695 (1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=86054470; PubMed=2824147;
RA Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Gall J.,
RA Hsiung N.;
RT "Expression of human uterine tissue-type plasminogen activator in
RT mouse cells using BBV vectors.";
RL DNA 6:461-472 (1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=86196143; PubMed=3009482;
RA Friezen Degen S.J., Rajput B., Reich E.;
RT "The human tissue plasminogen activator gene.";
RL J. Biol. Chem. 261:6972-6985 (1986).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=84298137; PubMed=6089198;
RA Ny T., Elgh F., Lund B.;
RT "The structure of the human tissue-type plasminogen activator gene:
RT correlation of intron and exon structures to functional and
RT structural domains.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359 (1984).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=86284200; PubMed=3090401;
RA Harris T.J., Patel T., Marston F.A., Little S., Emage J.S.,
RA Opendenker G., Volckaert G., Rombauts W., Billiau A., Somer P.;
RT "Cloning of cDNA coding for human tissue-type plasminogen activator
RT and its expression in Escherichia coli.";
RL Mol. Biol. Med. 3:279-292 (1986).
RN [7]
RP SEQUENCE OF 212-361 FROM N.A.
RX MEDLINE=83169656; PubMed=6572897;
RA Edlund T., Ny T., Raahy M., Heden L.-O., Palm G., Holmgren E.,
RA Josephson S.;
RT "Isolation of cDNA sequences coding for a part of human tissue
RT plasminogen activator.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:345-352 (1983).
RN [8]
RP SEQUENCE OF 1-36 FROM N.A.
RX MEDLINE=85289338; PubMed=3161893;
RA Fisher R., Waller E.K., Grosel G., Thompson D., Tizard R.,

```


RA Schleuning W.-D.;
 RT "Isolation and characterization of the human tissue-type plasminogen
 RT activator structural gene including its 5' flanking region."
 RL J. Biol. Chem. 260:11223-11230(1985).
 RN [9]
 RP SEQUENCE OF 31-562 FROM N.A.
 RX MEDLINE=91291340; PubMed=1368681;
 RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;
 RT "Purification and characterization of tissue plasminogen activator
 RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells."
 RL Agric. Biol. Chem. 55:1225-1232(1991).
 RN [10]
 RP SEQUENCE OF 36-562.
 RX TISSUE-Melanoma;
 RA MEDLINE=9500468; PubMed=6433976;
 RA Pohl G., Kaelin-Lang A., Bergsdorf N., Wallen P., Joernvall H.;
 RT "Tissue plasminogen activator: peptide analyses confirm an indirectly
 RT derived amino acid sequence, identify the active site serine residue,
 RT establish glycosylation sites, and localize variant differences."
 RL Biochemistry 23:3701-3707(1984).
 RN [11]
 RP SEQUENCE OF 33-52 AND 311-330.
 RX TISSUE-Melanoma;
 RA MEDLINE=93209620; PubMed=6682760;
 RA Wallen P., Pohl G., Bergsdorf N., Raandy M., Ny T., Joernvall H.;
 RT "Purification and characterization of a melanoma cell plasminogen
 RT activator."
 RL Eur. J. Biochem. 132:681-686(1983).
 RN [12]
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).
 RX TISSUE=Umbilical vein;
 RA MEDLINE=90191219; PubMed=2107528;
 RA Siebert P.D., Fong K.;
 RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
 RT human endothelial cells."
 RL Nucleic Acids Res. 18:1086-1086(1990).
 RN [13]
 RP STRUCTURE OF CARBOHYDRATES.
 RX MEDLINE=90092112; PubMed=2513186;
 RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;
 RT "Carbohydrate structure of recombinant human uterine tissue
 RT plasminogen activator expressed in mouse epithelial cells."
 RL Eur. J. Biochem. 186:273-286(1989).
 RN [14]
 RP CARBOHYDRATE-LINKAGE SITE THR-36.
 RX MEDLINE=91159408; PubMed=1900431;
 RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;
 RT "Tissue plasminogen activator has an O-linked fucose attached to
 RT threonine-61 in the epidermal growth factor domain."
 RL Biochemistry 30:2311-2314(1991).
 RN [15]
 RP DISULFIDE BONDS IN KRINGLE 2.
 RX MEDLINE=91244765; PubMed=1645336;
 RA Vilhose C.J., Wilhelm O.G., Hassell T., Jaskunas S.R., Bang N.U.;
 RT "Disulfide pairing of the recombinant kringle-2 domain of tissue
 RT plasminogen activator produced in *Bacteriella coli*."
 RL J. Biol. Chem. 266:10070-10072(1991).
 RN [16]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=96200985; PubMed=8613982;
 RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
 RA Bode W.;
 RT "Type 2,3 A crystal structure of the catalytic domain of recombinant
 RT two-chain human tissue-type plasminogen activator."
 RL J. Mol. Biol. 258:117-135(1996).
 RN [17]
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=97449126; PubMed=9305622;
 RA Renatus M., Engh R.A., Stubbs W.T., Huber R., Fischer S., Kohnert U.,
 RA Bode W.;
 RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray
 RT crystal structure of single-chain human tPA."
 RL EMBO J. 16:4797-4805(1997).

RN [18]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.
 RX MEDLINE=92118803; PubMed=1310033;
 RA de Vos A., Ultsch M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,
 RA Westbrook M.L., Kosiakof A.A.;
 RT "Crystal structure of the kringle 2 domain of tissue plasminogen
 RT activator at 2.4-A resolution."
 RL Biochemistry 31:270-279(1992).
 RN [19]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=90122799; PubMed=2558718;
 RA Byeon I.-U.L., Kelley R.F., Llinas M.;
 RT "1H NMR structural characterization of a recombinant kringle 2 domain
 RT from human tissue-type plasminogen activator."
 RL Biochemistry 28:9350-9360(1989).
 RN [20]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=92106329; PubMed=1762144;
 RA Byeon I.-U.L., Llinas M.;
 RT "Solution structure of the tissue-type plasminogen activator. 1H-NMR
 RT assignments and secondary structure."
 RL Eur. J. Biochem. 197:155-165(1991).
 RN [21]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=92106329; PubMed=1762144;
 RA Byeon I.-U.L., Llinas M.;
 RT "Solution structure of the tissue-type plasminogen activator. kringle
 RT 2 domain complexed to 6-aminohexanoic acid an antifibrinolytic
 RT drug."
 RL J. Mol. Biol. 222:1035-1051(1991).
 RN [22]
 RP STRUCTURE BY NMR OF 38-85.
 RX MEDLINE=92292163; PubMed=1602484;
 RA Downing A.K., Driscoll P.C., Harvey T.S., Dudgeon T.J., Smith B.O.,
 RA Baron M., Campbell I.D.;
 RT "The solution structure and backbone dynamics of the fibronectin type
 RT I and epidermal growth factor-like pair of modules of tissue-type
 RT plasminogen activator."
 RL J. Mol. Biol. 225:821-833(1992).
 RN [23]
 RP STRUCTURE BY NMR OF 36-126.
 RX MEDLINE=96027104; PubMed=7582899;
 RA Smith B.O., Downing A.K., Driscoll P.C., Dudgeon T.J., Campbell I.D.;
 RT "The solution structure and backbone dynamics of the fibronectin type
 RT I and epidermal growth factor-like pair of modules of tissue-type
 RT plasminogen activator."
 RL Structure 3:823-833(1995).
 CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMAGEN PLASMINOGEN
 CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
 CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
 CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
 CC MANY OTHER PHYSIOLOGICAL EVENTS.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: BINDS TO FIBRIN WITH HIGH AFFINITY. THIS INTERACTION

Query Match 11.1%; Score 154.5; DB 1; Length 562;
 Best Local Similarity 39.6%; Pred. No. 1.1e-05;
 Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
 QY 24 CFWDNGLHYREDQTSAPRLCLNWLDAQSLASAPVS-----GAGNSHYCNPDE 74
 DB 127 CYDQGISYRGWTSVAESGAECTNW--NSSALAQKPYSGRRPDAIRLIGLNHYCNPDR 184
 QY 75 DPRGPCVYSGEAGVBEKRPCEDLRCPETTS 105
 DB 185 DSK-PWCYVF-KAKGKSSSEFCSTPACSEGS 213

RESULT 3
 TPA RAT STRAND: PRT; 559 AA.
 ID TPA RAT
 AC P19637;

01-FEB-1991 (Rel. 17, Created)
01-NOV-1991 (Rel. 20, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
Tissue-type plasminogen activator precursor (BC 3.4.21.68) (CPA)
(t-PA) (t-plasminogen activator).
PLAT.
Rattus norvegicus (Rat).
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_taxid=1016;
[1]
SEQUENCE FROM N.A.
MEDLINE=8917014; PubMed=318445;
Ny T., Leonardsson G., Haseh A.J.W.;
"Cloning and characterization of a cDNA for rat tissue-type
plasminogen activator";
DNA 7:671-677(1988).
[2]
SEQUENCE FROM N.A.
MEDLINE=9010448; PubMed=2105315;
Feng P., Ohlsson M., Ny T.;
The structure of the tPA-less rat tissue-type plasminogen activator
gene. Species-specific sequence divergences in the promoter predict
differences in regulation of gene expression".
J. Biol. Chem. 265:2022-2027(1990).
-1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOMEN PLASMINOGEN
TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CONTROLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
MANY OTHER PHYSIOLOGICAL EVENTS.
-1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
plasminogen to form plasmin.
-1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
BOND.
-1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
-1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
-1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
-1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
-1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE 1 DOMAIN.
-1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.ebi.ch/announce/>
or send an email to license@ebi.ebi.ch).

EMBL; M23697; AAA41812.1; -;
EMBL; M31197; AAA42261.1; -;
EMBL; M31185; AAA42261.1; JOINED.
EMBL; M31186; AAA42261.1; JOINED.
EMBL; M31187; AAA42261.1; JOINED.
EMBL; M31188; AAA42261.1; JOINED.
EMBL; M31189; AAA42261.1; JOINED.
EMBL; M31190; AAA42261.1; JOINED.
EMBL; M31191; AAA42261.1; JOINED.
EMBL; M31192; AAA42261.1; JOINED.
EMBL; M31193; AAA42261.1; JOINED.
EMBL; M31194; AAA42261.1; JOINED.
EMBL; M31195; AAA42261.1; JOINED.
EMBL; M31196; AAA42261.1; JOINED.
EMBL; A19618; CA01482.1; -;
PIR; A31597; A31597.
PIR; A35029; A35029.
HSSP; P00750; 1RTE.
MEROPS; S01.232; -;
InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000083; Fibrinctn.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00039; fn1_1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF_1.
DR SMART; SM00058; fn1_1.
DR SMART; SM00130; KR_2.
DR SMART; SM00020; Tryp_Spc_1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN 1; 1.
DR PROSITE; PS00021; KRINGLE 2; 2.
DR PROSITE; PS50070; KRINGLE 2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM_1.
DR PROSITE; PS00134; TRYPSIN_HIS_1.
DR PROSITE; PS00135; TRYPSIN_SER_1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17
FT PROPEP 18 29
FT CHAIN 30 559
FT CHAIN 30 308
FT CHAIN 309 559
FT DOMAIN 36 78
FT DOMAIN 79 117
FT DOMAIN 124 205
FT DOMAIN 213 294
FT DOMAIN 309 559
FT ACT SITE 355 355
FT ACT SITE 404 404
FT ACT SITE 510 510
FT ACT SITE 38 68
FT DISULFD 66 75
FT DISULFD 83 94
FT DISULFD 88 105
FT DISULFD 107 116
FT DISULFD 124 205
FT DISULFD 145 187
FT DISULFD 176 200
FT DISULFD 213 294
FT DISULFD 234 276
FT DISULFD 265 289
FT DISULFD 297 428
FT DISULFD 340 356
FT DISULFD 348 417
FT DISULFD 442 516
FT DISULFD 474 490
FT DISULFD 506 534
FT CARBOHYD 149 149
FT CARBOHYD 481 481
FT CONFLICT 380 380
SQ SEQUENCE 559 AA; 62903 MW; 7DBD3809CD1DC921 CRC64;
Query Match 11.1%; Score 154; DB 1; Length 559;
Best Local Similarity 32.4%; Pred. No. 1.2e-05;
Matches 45; Conservative 14; Mismatches 54; Indels 26; Gaps 6;
DB 24 CFWNGHLYREDTSPAPRGRCNMLDAGSLAPVS-----GAGNSYGRNDE 74
DB 124 CREGGRTTRGTSTANENGABECTNM--NSSALSKRPYSARRPNVAKLGANHYCRNPDR 181
QY 75 DFRGFWCYVSGEAGVPEKRPCEDLRCP-----TTSQALPAFTTEIQEASGPG 123
DB 182 DVK-FWCYVF-KAGKYTEFCSTPACPKGPTEDCYVGKGVYRGTHSFTT--SKASCLPW 237

QY 124 ADEVQFAPANALPARESEA 142
Db 238 NSMILICKTAMRANSQA 256

RESULT 4
HGFA_MOUSE STANDARD; PRT; 653 AA.
AC 09R058: 09JRV4: 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update) (BC 3.4.21.-) (HGFA_MOUSE)
DE Hepatocyte growth factor activator precursor (BC 3.4.21.-) (HGFA_MOUSE)
GN HGFA_MOUSE
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OK NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Itoh H., Kataoka H., Koono H.;
RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA van Aelbeberg J.S., Sehgal S., Kukes A., Brady C., Barasch J., Yang J., Huan Y.;
RT "Activation of HGF by endogenous HGF activator is required for mesangial kidney morphogenesis in vitro."
RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM (BY SIMILARITY).
CC -1- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by, and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF099017; AAF02489.1; -
DR EMBL: AF224724; AAF34712.1; -
DR HSSP: P00763; IDPO.
DR MEROPS: S01.228; -
DR MGD: MGI:185281; Hgfac.
DR InterPro: IPR001314; Cytomotrypsin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF-2.
DR InterPro: IPR000562; FN_Type_II.
DR InterPro: IPR000083; Fibrinchnl.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_Protease_Try.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00038; fn1; 1.
DR Pfam: PF00040; fn2; 1.
DR Pfam: PF00051; kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00013; FNTYPEII.

DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 1.
DR ProDom: PD000995; FN_Type_II; 1.
DR SMART: SM00181; EGF_2.
DR SMART: SM0058; FN1; 1.
DR SMART: SM0059; FN2; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; TRYD_SPC; 1.
DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS01253; FIBRONECTIN_1; 1.
DR PROSITE: PS00023; FIBRONECTIN_2; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
DR PROSITE: PS050240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; 1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
KM Hydroxylase; Glycoprotein; Plasma; Serine protease; Kringle; signal; EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 29
FT PROPEP 30 369
FT CHAIN 370 405
FT FT CHAIN 406 653
FT FT DOMAIN 105 145
FT FT DOMAIN 157 195
FT FT DOMAIN 197 237
FT FT DOMAIN 238 276
FT FT DOMAIN 283 364
FT FT DOMAIN 406 653
FT FT ACT SITE 445 445
FT FT ACT SITE 495 495
FT FT ACT SITE 596 596
FT FT DISULFID 105 130
FT FT DISULFID 119 145
FT FT DISULFID 161 172
FT FT DISULFID 166 183
FT FT DISULFID 185 194
FT FT DISULFID 199 227
FT FT DISULFID 225 234
FT FT DISULFID 242 253
FT FT DISULFID 247 264
FT FT DISULFID 266 275
FT FT DISULFID 283 364
FT FT DISULFID 304 346
FT FT DISULFID 335 359
FT FT DISULFID 392 519
FT FT DISULFID 430 446
FT FT DISULFID 438 508
FT FT DISULFID 533 602
FT FT DISULFID 565 581
FT FT DISULFID 592 620
FT FT CARBOHYD 39 39
FT FT CARBOHYD 47 47
FT FT CARBOHYD 63 63
FT FT CARBOHYD 287 287
FT FT CARBOHYD 466 466
FT FT CARBOHYD 544 544
FT FT CONFLICT 164 164
SQ SEQUENCE 653 AA; 70567 MW; 88B4B2055DF7DC CRC64;

Query Match 10.8%; Score 150; DB 1; Length 653;
Best Local Similarity 32.6%; Pred. No. 3; 1e-05;
Matches 46; Conservative 17; Mismatches 46; Indels 32; Gaps 7;

QY 24 CFWDNGHYREDQSPAPGLRCINW-----LDQSGLASPVSGAGNHSYCNRPED 75
Db 283 CFLGNGTEYRGVASTAASGLSLMNSDLYOBLHYDS-VAAVVLGLSPHAYCNPED 341
QY 76 PRGFWCVVSGEAGVPEKPCEDLRCPEFTSQALPAFTTEIQEASBPGADEVQFAPANA 135
Db 342 ER-PWCYVVKDNLMSWE-----YCRLTACESLARVHSQTPPE-----ILA--A 380

QY 136 LPARSEAAVPCISQVR 156
DB 381 LP-ESAPAVPTCGKHKR 399

RESULT 5
TPA_MOUSE STANDARD; PRT; 559 AA.

AC P1214;
DT 01-JUN-1989 (Rel. 11, Created)
DT 01-JUN-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8808703; PubMed=2826484;
RA Rieckes R.J., Darow A.L., Strickland S.;
RT "Molecular cloning of complementary DNA to mouse tissue plasminogen
RT activator mRNA and its expression during F9 teratocarcinoma cell
RT differentiation.";
RL J. Biol. Chem. 263:1563-1569(1988).
CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: J03520; AAA0470.1; -.
DR PIR: A29941; A29941.
DR HSSP: P00750; IASH.
DR MEROPS: S01.232; -.
DR MGD: MGI:97610; PLAT.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000083; Fibrinctn.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00039; fn1; 1.
DR Pfam: PF00051; kringle; 2.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 2.

DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN 1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17 PROBABLE.
FT PROPEP 18 29
FT CHAIN 30 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 30 308 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT CHAIN 309 559 CHAIN.
FT CHAIN 309 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT DOMAIN 36 78 FIBRONECTIN TYPE-1.
FT DOMAIN 79 117 EGF-LIKE.
FT DOMAIN 124 205 KRINGLE 1.
FT DOMAIN 213 294 KRINGLE 2.
FT DOMAIN 309 559 SERINE PROTEASE.
FT ACT_SITE 355 355 CHARGE RELAY SYSTEM.
FT ACT_SITE 404 404 CHARGE RELAY SYSTEM.
FT ACT_SITE 510 510 CHARGE RELAY SYSTEM.
FT DISULFID 38 68 BY SIMILARITY.
FT DISULFID 66 75 BY SIMILARITY.
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 88 105 BY SIMILARITY.
FT DISULFID 107 116 BY SIMILARITY.
FT DISULFID 124 205 BY SIMILARITY.
FT DISULFID 145 187 BY SIMILARITY.
FT DISULFID 176 200 BY SIMILARITY.
FT DISULFID 213 294 BY SIMILARITY.
FT DISULFID 234 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 297 428 INTERCHAIN (BY SIMILARITY).
FT DISULFID 340 356 BY SIMILARITY.
FT DISULFID 348 417 BY SIMILARITY.
FT DISULFID 442 516 BY SIMILARITY.
FT DISULFID 474 490 BY SIMILARITY.
FT DISULFID 506 534 BY SIMILARITY.
FT CARBOHYD 149 149 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 559 AA; 63110 MW; 4ACE57DCA282A5 CRC64;

Query Match 10.6%; Score 147.5; DB 1; Length 559;
Best Local Similarity 37.0%; Pred. No. 4, 3e-05;
Matches 34; Conservative 11; Mismatches 34; Indels 13; Gaps 4;

QY 24 CPWDNGHYRSDQSPAPGLRCNWLDAOSGLASAVS-----GAGNHSYGRNDE 74
DB 124 CREEGQITRGSTWSTASGACLIW--NSSVLKPKYNAKRRPAIKLGLGNHNYCRNPD 181

QY 75 DRRGFWCVSGEAGVPEKRPCEDLRCPEITTSQ 106
DB 182 DLK-PWCYVF-KAGKTYTTECFSPACKGKSE 211

RESULT 6
HGFA_HUMAN STANDARD; PRT; 655 AA.
AC Q04756; Q14726;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
DE activator) (HGFA).
GN HGFAc.

OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RC TISSUE=Liver, and Serum; PubMed=763665;
 RX MEDLINE=93252878; PubMed=763665;
 RA Miyazawa K., Shimomura T., Kltamura A., Kondo J., Morimoto Y.,
 RA Kltamura N.;
 RT "Molecular cloning and sequence analysis of the cDNA for a human
 RT serine protease responsible for activation of hepatocyte growth
 RT factor. Structural similarity of the protease precursor to blood
 RT coagulation factor XII.";
 RT J. Biol. Chem. 268:10024-10028(1993).
 RN [2]
 RP SEQUENCE OF 40-655 FROM N.A.
 RA Zhao S., Odell C.;
 RL Submitted (Feb-1996) to the EMBL/GenBank/DBJ databases.
 CC -1 FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY
 CC CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM.
 CC -1 SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
 CC DISULFIDE BOND.
 CC -1 SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
 CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM.
 CC -1 TISSUE SPECIFICITY: LIVER.
 CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1 SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -1 SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
 CC -1 SIMILARITY: CONTAINS 1 KRINGEIN TYPE II DOMAIN.
 CC -1 SIMILARITY: CONTAINS 1 KRINGEIN DOMAIN.
 CC -1 CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D14012; BA03113.1; -;
 DR EMBL; Z69923; CA93803.1; -;
 DR PIR; A46688; A46688.
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.228; -;
 DR Genew; HGNC:4894; HGFAC.
 DR MIM; 604552; -;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000742; EGF 2.
 DR InterPro; IPR000562; FN Type II.
 DR InterPro; IPR000083; Fibrinctn.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00013; CHYMOTRYPSIN.
 DR PRINTS; PR00013; FNTPR1.
 DR PRINTS; PR00018; KRINGLE.
 DR PRODOM; PD000395; Kringle; 1.
 DR PRODOM; PD000395; FN_Type_II; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.

DR PROSITE; PS00023; FIBRONECTIN 2; 1.
 DR PROSITE; PS00021; KRINGLE 1; 1.
 DR PROSITE; PS00070; KRINGLE 2; 1.
 DR PROSITE; PS00240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 KW Hydrolyase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
 KW EGF-like domain; Repeat; Zymogen.
 FT SIGNAL 1 30
 FT PROPEP 31 372
 FT CHAIN 373 407
 FT CHAIN 408 655
 FT DOMAIN 108 148
 FT DOMAIN 160 198
 FT DOMAIN 200 240
 FT DOMAIN 241 279
 FT DOMAIN 286 367
 FT DOMAIN 408 655
 FT ACT_SITE 447 447
 FT ACT_SITE 497 497
 FT ACT_SITE 598 598
 FT ACT_SITE 108 133
 FT DISULFID 122 148
 FT DISULFID 154 175
 FT DISULFID 169 186
 FT DISULFID 188 197
 FT DISULFID 202 230
 FT DISULFID 228 237
 FT DISULFID 245 256
 FT DISULFID 250 267
 FT DISULFID 269 278
 FT DISULFID 286 367
 FT DISULFID 307 349
 FT DISULFID 338 362
 FT DISULFID 394 521
 FT DISULFID 432 448
 FT DISULFID 440 510
 FT DISULFID 535 604
 FT DISULFID 567 583
 FT DISULFID 594 622
 FT CARBOHYD 48 48
 FT CARBOHYD 290 290
 FT CARBOHYD 468 468
 FT CARBOHYD 492 492
 FT CARBOHYD 546 546
 FT CONFLICT 644 644
 SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1E1B862ED7 CRC64;
 Query Match 10.6%; Score 146.5; DB 1; Length 655;
 Best Local Similarity 36.9%; Pred. No. 6.2e-05;
 Matches 41; Conservative 10; Mismatches 41; Indels 19; Gaps 6;
 QY 24 CFMDNGHLYREDQTSPPRGRCINW-----LDAGSLASPVSGAGNHSYCRNPDED 75
 DB 286 CFLNGGTGYGVASTSAGSLCAANSDLIYQELHYDS-VGAALALGLPHACRPPDND 344
 QY 76 PRGMCVYSGEAGVP-----EKRPCEDLRCPEITTSQALPATTETIOE-ASEG 121
 DB 345 ER-FMCTIVVDLSALSWEXCLNACESL-----TRVQSPDLATLPSPAPSG 390
 RESULT 7
 ID PLNM HUMAN STANDARD; PRT; 810 AA.
 AC P00747;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
 GN FIG.
 OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90202879; PubMed=2318848;
RA Petersen T.E., Marzen M.R., Ichinose A., Davie E.W.;
RT "Characterization of the gene for human plasminogen, a key proenzyme
RL in the fibrinolytic system."; *J. Biol. Chem.* 265:6104-6111(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87162490; PubMed=3030813;
RA Forsgren M., Raden B., Israelsson M., Larsson K., Heden L.-O.;
RT "Molecular cloning and characterization of a full-length cDNA clone
RL for human plasminogen."; *FEBS Lett.* 213:254-260(1987).
RN [3]
RP SEQUENCE OF 20-810.
RA Sottrup-Jensen L., Petersen T.E., Magnusson S.;
RL Submitted (JUL-1977) to the EMBL data bank.
RN [4]
RP SEQUENCE OF 292-810 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
RL human and bovine plasminogen."; *Biochemistry* 23:4243-4250(1984).
RN [5]
RP SEQUENCE OF 20-100.
RX MEDLINE=75093329; PubMed=122932;
RA Wiman B., Wallen P.;
RT "Structural relationship between 'glutamic acid' and 'lysine' forms
RL of human plasminogen and their interaction with the NH₂-terminal
RT activation peptide as studied by affinity chromatography."; *Bur. J. Biochem.* 50:489-494(1975).
RN [6]
RP SEQUENCE OF 95-580; 581-626; 657-700 AND 732-810.
RA Sottrup-Jensen L., Claeys H., Zajdel M., Petersen T.E., Magnusson S.;
RL (in) Davidson J.F., Roman R.M., Samama M.M., Desnoyers P.C. (eds.);
RT Progress in chemical fibrinolysis and thrombolysis, pp.3:191-209,
RL Raven Press, New York (1978).
RN [7]
RP SEQUENCE OF 483-604.
RX MEDLINE=76043692; PubMed=126863;
RA Wiman B., Wallen P.;
RT "Amino-acid sequence of the cyanogen-bromide fragment from human
RL plasminogen that forms the linkage between the plasmin chains."; *Bur. J. Biochem.* 58:539-547(1975).
RN [8]
RP SEQUENCE OF 581-810.
RX MEDLINE=77252245; PubMed=142009;
RA Wiman B.;
RT "Primary structure of the B-chain of human plasmin."; *Bur. J. Biochem.* 76:129-137(1977).
RN [9]
RP ACTIVE SITE.
RX MEDLINE=73149248; PubMed=4694729;
RA Robbins K.C., Bernabe P., Arzadon L., Summaria L.;
RT "The primary structure of human plasminogen. II. The histidine loop
RL of human plasmin: light (B) chain active center histidine sequence."; *J. Biol. Chem.* 248:1631-1633(1973).
RN [10]
RP ACTIVE SITE.
RX MEDLINE=69234739; PubMed=4240117;
RA Groskopf W.R., Summaria L., Robbins K.C.;
RT "Studies on the active center of human plasmin. Partial amino acid
RL sequence of a peptide containing the active center serine residue."; *J. Biol. Chem.* 244:3590-3597(1969).
RN [11]
RP OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.
RX MEDLINE=82213905; PubMed=6919539;
RA Trexler M., Vail Z., Patchy L.;
RT "Structure of the omega-aminocarboxylic acid-binding sites of human
RL plasminogen. Arginine 70 and aspartic acid 56 are essential for
RT binding of ligand by kringle 4."; *J. Biol. Chem.* 257:7401-7406(1982).
RN [12]
RP FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.
RX MEDLINE=85054794; PubMed=6094526;
RA Vail Z., Patchy L.;
RT "The fibrin-binding site of human plasminogen. Arginines 32 and 34
RL are essential for fibrin affinity of the kringle 1 domain."; *J. Biol. Chem.* 259:13690-13694(1984).
RN [13]
RP PHOSPHORYLATION SITE SER-597.
RX MEDLINE=97345939; PubMed=9201958;
RA Wang H., Protok M., Bretthauer R.K., Castellino F.J.;
RT "Serine-578 is a major phosphorylation locus in human plasma
RL plasminogen."; *Biochemistry* 36:8100-8106(1997).
RN [14]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=88185329; PubMed=3356193;
RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamberling J.P.,
RA Gerwig G.J., Van Halbeek H., Vliegenthart J.F.;
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
RL plasminogen. Species specificity in relation to sialylation and
RT fucosylation patterns."; *Bur. J. Biochem.* 173:57-63(1988).
RN [15]
RP CARBOHYDRATE-LINKAGE SITE SER-268.
RX MEDLINE=97207306; PubMed=9054441;
RA Pitzer-Shepherd S.R., Stevens R.D., Andon N.L., Enghild J.J.,
RA Pizzo S.V.;
RT "Evidence for a novel O-linked sialylated triaccharide on Ser-248 of
RL human plasminogen 2."; *J. Biol. Chem.* 272:7408-7411(1997).
RN [16]
RP CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
RX MEDLINE=95042728; PubMed=7525077;
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
RL suppression of metastases by a Lewis lung carcinoma."; *Cell* 79:315-328(1994).
RN [17]
RP CHARACTERIZATION OF ANGIOSTATIN.
RX MEDLINE=97238710; PubMed=9102221;
RA Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Madsen J.W.,
RA Lapcevich R., Nacy C.A.;
RT "A recombinant human angiotensin protein inhibits experimental primary
RL and metastatic cancer."; *Cancer Res.* 57:1329-1334(1997).
RN [18]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.
RX MEDLINE=92031502; PubMed=1657148;
RA Mulichak A.M., Tulinsky A., Ravichandran K.G.;
RT "Crystal and molecular structure of human plasminogen kringle 4
RL refined at 1.9-A resolution."; *Biochemistry* 30:10576-10588(1991).
RN [19]
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.
RX MEDLINE=92031503; PubMed=1657149;
RA Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;
RT "The refined structure of the epsilon-aminocaproic acid complex of
RL human plasminogen kringle 4."; *Biochemistry* 30:10589-10594(1991).
RN [20]
RP X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
RX Stec B., Yamano A., Whitlow M., Teeter M.M.;
RT "Structure of human plasminogen kringle 4 at 1.68 Angstrom and 277 K.
RL Acta Crystallogr. D 53:169-178(1997).
RN [21]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.

```

RX MEDLINE=6180681; PubMed=611560;
RA Mathews I.I., Vanderhoff-Hanover P., Castellino F.J., Tulinsky A.;
RT "Crystal structures of the recombinant kringle 1 domain of human
RT plasminogen in complexes with the ligands epsilon-aminocaproic acid
RT and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";
RL Biochemistry 35:2567-2576(1996).
RN [22]
RP X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.
RX MEDLINE=98198034; PubMed=9521645;
RA Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,
RA Castellino F.J.;
RT "Structure and ligand binding determinants of the recombinant kringle
RT 5 domain of human plasminogen.";
RL Biochemistry 37:3258-3271(1998).
RN [23]
RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE=94237157; PubMed=8181475;
RA Rejzante M.R., Llinas M.;
RT "1H-NMR assignments and secondary structure of human plasminogen
RT kringle 1.";
RL Eur. J. Biochem. 221:927-937(1994).
RN [24]
RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE=94237158; PubMed=8181476;
RA Rejzante M.R., Llinas M.;
RT "Solution structure of the epsilon-aminohexanoic acid complex of
RT human plasminogen kringle 1.";
RL Eur. J. Biochem. 221:939-949(1994).
RN [25]
RP STRUCTURE BY NMR OF 183-354.
RX MEDLINE=96194155; PubMed=8652577;
RA Seehndel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,
RA Rickli E.E.;
RT "Recombinant gene expression and 1H NMR characteristics of the
RT kringle (2 + 3) supermodule: spectroscopic/functional individuality
RT of plasminogen kringle domains.";
RL Biochemistry 35:2357-2364(1996).
RN [26]
RP STRUCTURE BY NMR OF 374-461.
RX MEDLINE=90219023; PubMed=2157850;
RA Actinon R.A., Williams R.J.P.;
RT "Solution structure of the kringle 4 domain from human plasminogen by
RT 1H nuclear magnetic resonance spectroscopy and distance geometry.";
RL J. Mol. Biol. 212:541-552(1990).
RN [27]
RP VARIANTS PHE-374 AND THR-620.

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91250378; PubMed=1645711;
RA Kamalas J.J., Makher S.P.;
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT receptor site for plasminogen.";
RL J. Biol. Chem. 266:10825-10829(1991).
RN [2]
RP FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
RT A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
RT EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
RT AND INFLAMMATION. IT WEAKENS THE WALLS OF THE
RT GRAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
RT ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
RT AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
RT LAMININ AND VON WILLEBRAND FACTOR.
RT CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
RT higher selectivity than trypsin. Converts fibrin into soluble
RT products.
RN [3]
RP ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
RT ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
RT FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
RN [4]
RP MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
RT IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
RN [5]
RP SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
RN [6]
RP THIS SWISS-PROT entry is copyright. It is produced through a collaboration
RT between the Swiss Institute of Bioinformatics and the EMBL outstation -
RT the European Bioinformatics Institute. There are no restrictions on its
RT use by non-profit institutions as long as its content is in no way
RT modified and this statement is not removed. Usage by and for commercial
RT entities requires a license agreement (See http://www.isb-sib.ch/announce/
RT or send an email to license@sib-sib.ch).
RN [7]
RP EMBL: M62832; AAA41884.1; -.
RN [8]
RP PIR: A40522; A40522.
RN [9]
RP HSSP: P00747; 1PMK.
RN [10]
RP MEROPS: S01.233; -.
RN [11]
RP InterPro: IPR000001; Kringle.
RN [12]
RP InterPro: IPR001254; Ser_protease_Try.
RN [13]
RP Pfam: PF00051; Kringle; 2.
RN [14]
RP ProDom: PD000395; Kringle; 2.
RN [15]
RP SMART: SM00130; KR; 1.
RN [16]
RP PROSITE: PS00021; KRINGLE 1; 1.
RN [17]
RP PROSITE: PS00070; KRINGLE 2; 2.
RN [18]
RP PROSITE: PS00240; TRYPSIN_DOM; PARTIAL.
RN [19]
RP PROSITE: PS00134; TRYPSIN_HIS; PARTIAL.
RN [20]
RP PROSITE: PS00135; TRYPSIN_SER; PARTIAL.
RN [21]
RP HydroLase: Serine protease; Plasma; Glycoprotein; Fibrinolysis;
RN [22]
RP Tissue remodeling; Blood coagulation; Kringle; Repeat.
RN [23]
RP NON_TER 1 1
RN [24]
RP DOMAIN 1 10
RN [25]
RP FT <1 10 KRINGLE 3 (BY SIMILARITY).
RN [26]
RP FT 34 112 KRINGLE 4 (BY SIMILARITY).
RN [27]
RP FT 139 >169 KRINGLE 5 (BY SIMILARITY).
RN [28]
RP FT DISULFD 34 112 BY SIMILARITY.
RN [29]
RP FT DISULFD 55 95 BY SIMILARITY.
RN [30]
RP FT DISULFD 83 107 BY SIMILARITY.
RN [31]
RP FT NON_TER 169 169
RN [32]
RP SEQUENCE 169 AA; 18401 MW; 77A54214C49D010C CRC64;

```

Query Match 9.9%; Score 138; DB 1; Length 169;
Best Local Similarity 31.8%; Pred. No. 7.2e-05;
Matches 35; Conservative 13; Mismatches 48; Indels 14; Gaps 5;

```

QY 24 CFWNNGHLYREDQTSBAPGLRCLMWLDA-----QGLASAPVSGAGNHSYCRNPDEPRG 78
DB 34 CYQNGSKSYRGTSTTNGKCKQSWMTSPHSHSKTPANFPDSG-EMNYCRNPDPDRG 92
QY 79 PWCYVSGAGVPEKR--PCEDLRCPETTSQALPFTTQGEASBGRGDE 126
DB 93 PWCFTTD---PSVMEYCNLKRCSGTGGV--AESALVPQVSPAPGRS 136

```


RESULT 9
UOK_PAPCY STANDARD; PRT; 433 AA.
AC P16227;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Papio cyncephalus (Yellow baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopitheciinae; Papio.
OX NCBI_TaxID=9556;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thoracic aorta;
RX MEDLINE=90287734; PubMed=2113276;
RA Au Y.P.T., Wang T.W., Clowes A.W.;
RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
RT plasminogen activator."
RL Nucleic Acids Res. 18:3411-3411(1990).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X51935; CAA36200.1; -.
DR PIR; S14687; UKBAY.
DR HSSP; P00749; ILMW.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activator; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 1 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 176 CHAIN A (BY SIMILARITY).
FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).
FT CHAIN 178 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 26 62 EGF-LIKE.
FT DOMAIN 69 150 KRINGLE.

FT DOMAIN 151 177 CONNECTING PEPTIDE.
FT DOMAIN 178 433 SERINE PROTEASE.
FT DISULFID 30 38 BY SIMILARITY.
FT DISULFID 32 50 BY SIMILARITY.
FT DISULFID 52 61 BY SIMILARITY.
FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).
FT DISULFID 208 224 BY SIMILARITY.
FT DISULFID 216 287 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT SITE 223 273 CHARGE RELAY SYSTEM.
FT ACT SITE 274 274 CHARGE RELAY SYSTEM.
FT ACT SITE 378 378 CHARGE RELAY SYSTEM.
FT CARBOHYD 324 324 N-LINKED (GLCNAC...) (BY SIMILARITY).
SQ SEQUENCE 433 AA; 48595 MW; 816D220FDDC8792 CRC64;

Query Match 9.9%; Score 138; DB 1; Length 433;
Best Local Similarity 32.7%; Pred. No. 0.00021;
Matches 32; Conservative 14; Mismatches 36; Indels 16; Gaps 4;

QY 24 CFWDNGHLYREDQTSAPPLRCINWIDA-----OSGLASAPVSGAGHSYCRNPDEDP 76
DB 69 CYBNGGHFYRGKASTDTMGKSCIAWNSATVLOQTYAHRSDALQLGLGKHYCRNPDP 127
QY 77 RGPWCYVSGEAGVPEK-----RPCEDLRCPETTSQAL 108
DB 128 RRPWCYV--OVGLKQVQECMVHNCADGKKPSSPEEL 163

RESULT 10
UOK_PIG STANDARD; PRT; 442 AA.
AC P04185;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=85087954; PubMed=6096832;
RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
RL Nucleic Acids Res. 12:9525-9541(1984).
RN [2]
RP REVISION TO 241.
RA Nagamine Y.;
RL Submitted (DEC-1986) to the PIR data bank.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X01648; CAA25806.1; -.
DR EMBL; X02724; CAA26511.1; -.
DR PIR; A00932; UKPG.
DR HSSP; P00749; IKDU.

DR MEROPS: S01.231; -
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000001; Kringle.
 DR Pfam: PF0001254; Ser_protease_Try.
 DR Pfam: PF00089; Kringle; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR ProDom: PD000395; Kringle; 1.
 DR SMART: SM00130; KR; 1.
 DR SMART: SM00020; TRYP_SPC; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01185; EGF_2; FALSE_NEG.
 DR PROSITE: PS00022; KRINGLE_1; 1.
 DR PROSITE: PS00070; KRINGLE_2; 1.
 DR PROSITE: PS50240; TRYPsin_DOM; 1.
 DR PROSITE: PS00135; TRYPsin_HIS; 1.
 DR PROSITE: PS00135; TRYPsin_SER; 1.
 DR PLASMINOGEN ACTIVATION; Hydrolyase; Serine protease; Glycoprotein;
 KW Kringle; EGF-like domain; Zymogen; Signal.
 FT SIGNAL 1
 FT CHAIN 21 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 21 188 CHAIN A (BY SIMILARITY).
 FT CHAIN 190 442 CHAIN B (BY SIMILARITY).
 FT DOMAIN 29 65 EGF-LIKE.
 FT DOMAIN 72 153 KRINGLE.
 FT DOMAIN 154 189 CONNECTING PEPTIDE.
 FT DOMAIN 190 442 SERINE PROTEASE.
 FT CARBOHYD 152 152 N-LINKED (GLYCNC. . .).
 FT DISULFID 33 41 BY SIMILARITY.
 FT DISULFID 35 53 BY SIMILARITY.
 FT DISULFID 55 64 BY SIMILARITY.
 FT DISULFID 179 310 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 220 236 BY SIMILARITY.
 FT DISULFID 228 299 BY SIMILARITY.
 FT DISULFID 324 393 BY SIMILARITY.
 FT DISULFID 356 372 BY SIMILARITY.
 FT DISULFID 383 411 BY SIMILARITY.
 FT ACT_SITE 235 235 CHARGE RELAY SYSTEM.
 FT ACT_SITE 286 286 CHARGE RELAY SYSTEM.
 FT ACT_SITE 387 387 CHARGE RELAY SYSTEM.
 FT ACT_SITE 241 241 Q -> H (IN REF. 1; CAA25806).
 FT CONFLICT 242 242 O -> H (IN REF. 1; CAA25811).
 FT CONFLICT 288 288 A -> GS (IN REF. 1; CAA25806).
 FT CONFLICT 288 288 A -> GS (IN REF. 1; CAA25806).
 FT SEQUENCE 442 AA; 49116 MW; EB32FCE501321EB CRC64;

Query_Match 9.8%; Score 135.5; DB 1; Length 442;
 Best Local Similarity 36.9%; Pred. No. 0.00034;
 Matches 38; Conservative 7; Mismatches 37; Indels 21; Gaps 5;

QY 24 CFWDNGHLVREDOTSPAPGLRCINMLDAOSGL---ASAPVS---GAGNHSYCHNPDEDP 76
 DB 72 CFEQNGHSHYRGKANTNGGRPCLPMSATVLTINTYAHHPDLAQGLGKHNVCNPD-NQ 130
 QY 77 RGPACVYS-----GEAGVP-----EKRPCEDLRCPESTSQ 106
 DB 131 RRPACVYGVGAKQLVGECWCPNCGSGESHRPAYDGNKPFSTPE 173

RESULT 11
 UNTRB_DESRO STANDARD; PRT; 431 AA.
 AC P98121;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA beta)
 OS Desmodus rotundus (Vampire bat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 OC Desmodontinae; Desmodus.
 NCBI_taxid=9430;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=salivary gland;
 RX MEDLINE=92039036; PubMed=1937019;
 RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
 RA Alagon A., Donner P., Schleuning W.D.;
 RT "The plasminogen activator family from the salivary gland of the
 RT vampire bat Desmodus rotundus: cloning and expression.";
 RL Gene 105:229-237(1991).
 RN 12]
 RP CHARACTERIZATION
 RX MEDLINE=93393059; PubMed=1309059;
 RA Schleuning W.D., Alagon A., Boidol W., Bringmann P., Petri T.,
 RA Kraetzschmar J., Haendler B., Langer G., Baldu B., Witt W.,
 RA Donner P.;
 RT "Plasminogen activators from the saliva of Desmodus rotundus (common
 RT vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
 CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
 CC EXCLUSIVELY HAEMOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
 CC AGENT.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M63989; AAA31594.1; -
 CC HSSP: P98119; IAS1.
 DR MEROPS: S01.239; -
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000001; Kringle.
 DR Pfam: PF00089; Kringle; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR ProDom: PD000395; Kringle; 1.
 DR SMART: SM00181; EGF; 1.
 DR SMART: SM00130; KR; 1.
 DR SMART: SM00020; TRYP_SPC; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS00070; KRINGLE_2; 1.
 DR PROSITE: PS50240; TRYPsin_DOM; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; 1.
 DR PROSITE: PS00135; TRYPsin_SER; 1.
 DR PLASMINOGEN ACTIVATION; Hydrolyase; Serine protease; Glycoprotein;
 KW Kringle; EGF-like domain; Signal; Multigene family.
 FT SIGNAL 1
 FT CHAIN 37 431 SALIVARY PLASMINOGEN ACTIVATOR BETA.
 FT DOMAIN 37 75 EGF-LIKE.
 FT DOMAIN 82 163 KRINGLE.
 FT DOMAIN 179 431 SERINE PROTEASE.
 FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 275 275 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 382 382 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 41 32 BY SIMILARITY.
 FT DISULFID 46 63 BY SIMILARITY.
 FT DISULFID 65 74 BY SIMILARITY.
 FT DISULFID 82 163 BY SIMILARITY.

```

FT DISULFID 103 145 BY SIMILARITY.
FT DISULFID 134 158 BY SIMILARITY.
FT DISULFID 168 229 BY SIMILARITY.
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 288 BY SIMILARITY.
FT DISULFID 313 388 BY SIMILARITY.
FT DISULFID 345 361 BY SIMILARITY.
FT DISULFID 378 406 BY SIMILARITY.
FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 431 AA; 48221 MW; 699856758162CBP CRC64;

Query Match 9.7%; Score 134; DB 1; Length 431;
Best Local Similarity 38.6%; Pred. No. 0.00045;
Matches 27; Conservative 9; Mismatches 20; Indels 14; Gaps 3;

QY 24 CFMDNCHLYREDOQTSPAPGRCLNWLDAOSGL-----ASAPVSGAGNHSYCRNP 73
DB 82 CTKDQGVYKGTWSTSESGAQCINM--NSNLTRTYNGRSDALITLGLGNHNYCRNP 138
QY 74 EDRPGPCVY 83
DB 139 NNSK-PMCVY 147

RESULT 12
UR22_DESRO STANDARD; PRT; 477 AA.
ID 15638;
AC 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.66) (DSFA
DE alpha-2) (BAC-PA) (1-plasminogen activator).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OC NCBI_TaxID=9430;
RN 11;
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=salivary gland;
RX MEDLINE=92039367; PubMed=1937019;
RA Kraetzschmar U., Haendler B., Langer G., Bojdo W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RT Gene 105:229-237(1991).
RN 12;
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=salivary gland;
RX MEDLINE=90036867; PubMed=2509450;
RA Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,
RA Regester R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;
RT "Isolation, characterization, and cDNA cloning of a vampire bat
RT salivary plasminogen activator.";
RT J Biol. Chem. 264:17947-17952(1989).
RN 13;
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Bojdo W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RT Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC AGENT.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN
CC THE PRESENCE OF FIBRIN I.
CC

CC -1- SUBUNIT: MONOMER.
CC -1- DOMAIN: THE FIBRONECTIN TYPE-1 DOMAIN MEDIATES BINDING TO FIBRIN,
CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
CC STIMULATION OF ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE 1 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to licences@isb-sib.ch).
CC
CC EMBL, M63988; AAA31593.1; -.
CC EMBL, J05082; AAA31596.1; -.
CC PIR, A34369; A34369.
CC
CC HSSP, P98119; 1A51.
CC
CC MEROPS, S01.232; -.
CC
CC InterPro: IPR001314; Chymotrypsin.
CC InterPro: IPR000561; EGF-like.
CC InterPro: IPR000083; Fibrinectn.
CC InterPro: IPR000001; Kringle.
CC InterPro: IPR001254; Ser_protase_Try.
CC Pfam: PF00008; EGF_1.
CC Pfam: PF00039; fn1_1.
CC Pfam: PF00051; kringle_1.
CC Pfam: PF00085; trypsin_1.
CC PRINTS, PR00722; CHYMOTRYPSIN.
CC PRINTS, PR00018; KRINGLE.
CC ProDom, PD000395; Kringle_1.
CC SMART, SM00181; EGF_1.
CC SMART, SM00058; FN1_1.
CC SMART, SM00130; KR_1.
CC SMART, SM00020; Tryp_Spc_1.
CC PROSITE, PS00022; EGF_1; 1.
CC PROSITE, PS01166; EGF_2; 1.
CC PROSITE, PS01253; FIBRONECTIN_1; 1.
CC PROSITE, PS00021; KRINGLE_1; 1.
CC PROSITE, PS00020; KRINGLE_2; 1.
CC PROSITE, PS02400; TRYPsin_DOM; 1.
CC PROSITE, PS00134; TRYPsin_HIS; 1.
CC PROSITE, PS00135; TRYPsin_SER; 1.
CC KX Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
CC Kringle; EGF-like domain; Signal; Multigene family.
CC FT SIGNAL 1 36
CC FT CHAIN 37 477
CC FT DOMAIN 40 82
CC FT DOMAIN 83 121
CC FT DOMAIN 128 209
CC FT DOMAIN 225 477
CC FT ACT_SITE 272 272
CC FT ACT_SITE 321 321
CC FT ACT_SITE 428 428
CC FT DISULFID 42 72
CC FT DISULFID 70 79
CC FT DISULFID 87 98
CC FT DISULFID 92 109
CC FT DISULFID 111 120
CC FT DISULFID 128 209
CC FT DISULFID 149 191
CC FT DISULFID 180 204
CC FT DISULFID 214 345
CC FT DISULFID 237 273
CC FT DISULFID 265 334
CC FT DISULFID 359 434
CC FT DISULFID 391 407
CC FT DISULFID 424 452
CC FT CARBOHYD 185 185
CC FT CARBOHYD 398 398
CC
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CONFLICT 403 403 N -> K (IN REF. 2).
FT CONFLICT 417 417 Y -> H (IN REF. 2).
FT CONFLICT 435 435 M -> R (IN REF. 2).
SQ SEQUENCE 477 AA; 53719 MW; 17486555C0E5077C CRC64;

Query Match 9.7%; Score 134; DB 1; Length 477;
Best Local Similarity 38.6%; Pred. No. 0.0005;
Matches 27; Conservative 9; Mismatches 20; Indels 14; Gaps 3;

OY 24 CFMNGHLYREDQTSFAPGLRCLNWLDAOSGL-----ASAPVSGAGNHSYCRND 73
DB 128 CYKQGVYRGWTSSTSGAQCIW---NSNLTFRRTYNGRRSDA1TGLGNHYVCND 184
OY 74 EDPGPMCYV 83
DB 185 NNSK-PMCYV 193

RESULT 13
HGFL_MOUSE STANDARD; PRT; 716 AA.
ID HGFL_MOUSE
AC P26928;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor-like protein precursor (Macrophage
DE stimulatory protein) (MSP).
GN MST1 OR HGFL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10990;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/c; TISSUE=Liver;
RX MEDLINE=92002017; PubMed=1832957;
RA Fritzeneger Degen S.U., Stuart L.A., Han S., Jamison C.S.;
RT "Characterization of the mouse cDNA and gene coding for a hepatocyte
RT growth factor-like protein: expression during development."
RL Biochemistry 30:9781-9791(1991).
CC -1- FUNCTION: PROBABLY HAS NO PROTEOLYTIC ACTIVITY, SINCE CRUCIAL AA
CC CHARACTERISTIC OF SERINE PROTEASES CATALYTIC SITES ARE NOT
CC CONSERVED.
CC -1- TISSUE SPECIFICITY: LIVER. LOWER LEVELS IN LUNG, PLACENTA AND
CC ADRENAL.
CC -1- DEVELOPMENTAL STAGE: IS EXPRESSED AT LOW LEVELS DURING GESTATION.
CC JUST BEFORE BIRTH THE LEVEL INCREASES DRAMATICALLY AND REMAINS
CC STABLE AFTERWARDS.
CC -1- PMT. MAY BE CLEAVED AFTER AA 488, TO YIELD A TWO-CHAIN MOLECULE
CC HEUD TOGETHER BY DISULFIDE BONDS, OR CLEAVED INTO TWO SEPARATE
CC POLYPEPTIDES.
CC -1- SIMILARITY: BELONGS TO PERTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M74180; AAA50166.1; -
DR EMBL; M74181; AAA50167.1; -
DR HSSP; P00747; 1KRN.
DR MEROPS; S01.975; -
DR MGD; MGI:96080; HGFL.
DR InterPro; IPR001334; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00024; PAN; 1.

```

```

DR Pfam; PR00051; kringle_4.
DR Pfam; PR00089; trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_Ap; 1.
DR SMART; SM00020; TRY_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR Kringle; Glycoprotein; Serine protease homolog; Repeat; Signal.
KW SIGNAL;
FT CHAIN 1 716
FT DOMAIN 19 109
FT DOMAIN 110 186
FT DOMAIN 191 268
FT DOMAIN 292 370
FT DOMAIN 379 457
FT DOMAIN 489 716
FT DISULFID 56 78
FT DISULFID 60 66
FT DISULFID 110 186
FT DISULFID 131 169
FT DISULFID 157 181
FT DISULFID 191 268
FT DISULFID 194 333
FT DISULFID 212 251
FT DISULFID 240 263
FT DISULFID 292 370
FT DISULFID 313 352
FT DISULFID 341 364
FT DISULFID 379 457
FT DISULFID 400 440
FT DISULFID 428 452
FT DISULFID 477 593
FT DISULFID 512 528
FT DISULFID 607 672
FT DISULFID 637 651
FT DISULFID 651 651
FT CARBOHYD 72 72
FT CARBOHYD 173 173
FT CARBOHYD 305 305
FT CARBOHYD 620 620
FT CONFLICT 19 19
SQ SEQUENCE 716 AA; 80588 MW; BBCE02EF85213ACC CRC64;

Query Match 9.7%; Score 134; DB 1; Length 716;
Best Local Similarity 28.2%; Pred. No. 0.00079;
Matches 42; Conservative 9; Mismatches 46; Indels 52; Gaps 7;

OY 1 LLAVYQAFVSNMMLAEVAGSG-----GCPMDNGHLYREDQTSFAPGLRCLNM 48
DB 80 LHPYTO-----HSHTQVLSLSLCHLFQKQDYVRCIMDNQVSRGVARTAGLGLCGAM 134
OY 49 ---LDAOSGLASAPVSGAGNHSYCRNDEDPGPMCYV----- 84
DB 135 SRPNDHKTTPYKNG-ENFPCRNDDGPRSPWCYTNNSVRFQSGIKTCEAVCYL 193
OY 85 -----GEAGVPEK-RPCE--DIRCET 103
DB 194 CNGEDYGEVDVTEGRCRCORWDLQHPHS 222

RESULT 14
URTG_DESRO STANDARD; PRT; 394 AA.
ID URTG_DESRO
AC P49150;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator gamma precursor (BC 3.4.21.68) (DSPA
DE gamma).

```

OS Desmodus rotundus (Vampire bat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 OC Desmodontinae; Desmodus.
 NCBI_TaxID=9430;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=92039056; PubMed=1977019;
 RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
 RT "The plasminogen activator family from the salivary gland of the
 RL vampire bat Desmodus rotundus: cloning and expression.";
 RL Gene 105:229-237 (1991).
 RP CHARACTERIZATION.
 RX MEDLINE=93393059; PubMed=1309059;
 RA Schleming W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
 RA Kraetzschmar J., Haendler B., Langer G., Balduz B., Wilt W.,
 RT "Plasminogen activators from the saliva of Desmodus rotundus (common
 RT vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 667:395-403 (1992).
 CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
 CC EXCLUSIVELY HAEMOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
 CC AGENT.
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -!- SUBUNIT: MONOMER.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC -----
 DR EMBL; M63990; AAA31595.1; -.
 DR HSSP; P98119; IAS1.
 DR MEROPS; S01.239; -.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF000651; Kringle_1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRODOM; PD00018; KRINGLE.
 DR PRODOM; PD000395; Kringle_1.
 DR SMART; SM00130; KR_1.
 DR SMART; SM00020; TRYPSIN_1.
 DR PROSITE; PS00021; KRINGLE_2_1.
 DR PROSITE; PS50070; TRYPSIN_DOM_1.
 DR PROSITE; PS50240; TRYPSIN_HIS_1.
 DR PROSITE; PS00134; TRYPSIN_HIS_1.
 DR PROSITE; PS00135; TRYPSIN_SER_1.
 KW Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
 KW Kringle; Signal; Multigene family.
 FT SIGNAL 1
 FT CHAIN 36
 FT CHAIN 37
 FT DOMAIN 394
 FT DOMAIN 45
 FT DOMAIN 126
 FT ACT_SITE 142
 FT ACT_SITE 189
 FT ACT_SITE 238
 FT ACT_SITE 345
 FT ACT_SITE 345
 FT DISULFID 45
 FT DISULFID 126
 FT DISULFID 66
 FT DISULFID 97
 FT DISULFID 121
 FT DISULFID 131
 FT DISULFID 131
 FT DISULFID 174
 FT DISULFID 174
 FT DISULFID 182
 FT DISULFID 251

FT DISULFID 276
 FT DISULFID 308
 FT DISULFID 341
 FT CARBOHYD 315
 SQ SEQUENCE 394 AA; 44105 MW; 9CDD6F52F3D81FDC CRC64;
 Query Match 9.54; Score 132; DB 1; Length 394;
 Best Local Similarity 32.74; Pred. No. 0.0006;
 Matches 33; Conservative 9; Mismatches 39; Indels 20; Gaps 5;
 QY 16 AAYGSG--GCGFMDNGHLYXEDOTSPPAPGLRLTM-----LDAGSGLASAPVSGAGNH 66
 DB 35 SKAIGPAPATCKKQGGVTKQWTSSESQAQICINMNSMLIRRYNGRMPBAVKLGIDNH 94
 QY 67 SYCRNPDDEPRGPGVCY-----SGEAGYPERKPCEDLRC 100
 DB 95 NYCRNPDGASK-PWCYVTKARKRTSSCSVP--VCSKATC 131
 RESULT 15
 ID UROK_HUMAN STANDARD; PRT; 431 AA.
 AC P00749; Q15844; Q16618; Q969W6;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 20-MAR-1987 (Rel. 04, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85215647; PubMed=2987867;
 RA Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blasi F.;
 RT "The human urokinase-plasminogen activator gene and its promoter.";
 RL Nucleic Acids Res. 13:2759-2771 (1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,
 RA Steffens G.J., Heyneker H.L.;
 RT "Cloning and expression of the gene for pro-urokinase in Escherichia
 RT coli.";
 RL Biotechnology 3:923-929 (1985).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86056954; PubMed=2415429;
 RA Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Arimura H.,
 RA Niishida M., Suyama T.;
 RT "Molecular cloning of cDNA coding for human prepro-urokinase.";
 RL Gene 56:183-188 (1985).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85203359; PubMed=2888571;
 RA Jacobs P., Cravador A., Loriau R., Brockly F., Colau B., Chuchana P.,
 RA van Elsen A., Hertzog A., Bollen A.;
 RT "Molecular cloning, sequencing, and expression in Escherichia coli of
 RT human prepro-urokinase cDNA.";
 RL DNA 4:139-146 (1985).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Peol C.L., Yi Q.,
 RA Nickerson D.A.;
 RT Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA Strausberg R.;
 RT Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 66-431 FROM N.A.

RX MEDLINE=84272706; PubMed=6589620;
 RA Verde P., Scoppelli M.P., Galeffi P., di Nocera P., Biasi F.;
 RT "Identification and primary sequence of an unspliced human urokinase
 RT poly(A)+ RNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
 RN [8]
 RP SEQUENCE OF 21-177
 RA MEDLINE=83055084; PubMed=6754569;
 RX Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,
 RA Flohe L.;
 RT "The primary structure of high molecular mass urokinase from human
 RT urine. The complete amino acid sequence of the A chain.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
 RN [9]
 RP SEQUENCE OF 156-176 AND 179-224.
 RX MEDLINE=83003608; PubMed=6749491;
 RA Schaller J., Nick H., Rickli E.E., Gillissen D., Lergler W.,
 RA Studer R.O.;
 RT "Human low-molecular-weight urinary urokinase. Partial
 RT characterization and preliminary sequence data of the two polypeptide
 RT chains.";
 RL Eur. J. Biochem. 125:251-257(1982).
 RN [10]
 RP SEQUENCE OF 158-410.
 RX MEDLINE=83055099; PubMed=6754572;
 RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.;
 RT "The complete amino acid sequence of low molecular mass urokinase
 RT from human urine.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=96000858; PubMed=8591045;
 RA Spraggon C.M., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
 RA Dobson C.M., Stuart D.I., Jones E.Y.;
 RT "The crystal structure of the catalytic domain of human
 RT urokinase-type plasminogen activator.";
 RL Structure 3:681-691(1995).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
 RX MEDLINE=20266327; PubMed=10805774;
 RA Sperl S., Jacob U., Arroyo de Prada N., Stuerbecher J., Wilhelm O.G.,
 RA Boe W., Magdolen V., Huber R., Moroder L.;
 RT "(4-antimetabolite)phenylguanidine derivatives as nonpeptidic highly
 RT selective inhibitors of human urokinase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
 RN [13]
 RP STRUCTURE BY NMR.
 RX MEDLINE=89127526; PubMed=2535903;
 RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.;
 RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
 RT dimensional NMR.";
 RL Nature 337:579-582(1989).
 RN [14]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=93003110; PubMed=1327118;
 RA Li X., Smith R.A.G., Dobson C.M.;
 RT "Sequential IH NMR assignments and secondary structure of the kringles
 RT domain from urokinase.";
 RL Biochemistry 31:9562-9571(1992).
 RN [15]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=94149701; PubMed=8107093;
 RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.;
 RT "Solution structure of the kringles domain from urokinase-type
 RT plasminogen activator.";
 RL J. Mol. Biol. 235:1548-1559(1994).
 RN [16]
 RP VARIANT LEU-141.
 RX MEDLINE=96186279; PubMed=8652631;
 RA Yoshimoto M., Uehiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,
 RA Sasaki Y., Hanada K.;
 RT "Characterization of single chain urokinase-type plasminogen
 RT activator with a novel amino-acid substitution in the kringles

RT structure.";
 RL Biochim. Biophys. Acta 1293:83-89(1996).
 RN [17]
 RP VARIANT LEU-141.
 RX MEDLINE=97218551; PubMed=9065988;
 RA Conne B., Berczy M., Belin D.;
 RT "Detection of polymorphisms in the human urokinase-type plasminogen
 RT activator gene.";
 RL Thromb. Haemost. 77:434-435(1997).
 RN [18]
 RP ERRATUM.
 RA Conne B., Berczy M., Belin D.;
 RL Thromb. Haemost. 76:973-973(1997).
 RN [19]
 RP VARIANT LEU-141.
 RX MEDLINE=97337920; PubMed=9194591;
 RA Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Hell W.,
 RA Creutzburg S., Graeff H., Magdolen V.;
 RT "Mutation analysis of the genes encoding urokinase-type plasminogen
 RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
 RL Electrophoresis 18:686-689(1997).
 CC -1- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR
 CC THERAPY OF THROMBOLYTIC DISORDERS.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
 CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
 CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
 CC MOLECULAR MASS FORM TO YIELD A SHORT A CHAIN.
 CC -1- PHARMACEUTICAL: Available under the name Abbotkinase (Abbott). Used
 CC in Pulmonary Embolism (PE) to initiate fibrinolysis.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X02419; CAA26268.1; -;
 DR EMBL: M15476; AAA61253.1; -;
 DR EMBL: D00244; BAA00175.1; -;
 DR EMBL: D11143; BAA01919.1; -;
 DR EMBL: X02760; CAA26535.1; -;
 DR EMBL: A8377330; AAK53822.1; -;
 DR EMBL: BC013575; AAH13575.1; -;
 DR EMBL: K03226; AAC97138.1; -;
 DR EMBL: K02286; AAA61252.1; -;
 DR EMBL: A21571; CAA01559.1; -;
 DR EMBL: A18397; CAA01390.1; -;
 DR PIR: A00931; UKHU.
 DR PIR: A32974; A32974.
 DR PDB: 1KDU; 31-OCT-93.
 DR PDB: 1LMW; 29-JAN-96.
 DR PDB: 1URK; 08-MAY-95.
 DR PDB: 1EUN; 17-MAY-00.
 DR MEROPS: S01.231; -;
 DR GlycoSiteDB: P00749; -;
 DR Genew: HGNC:9052; PLAU.
 DR MIM: 191840; -;
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Ser protease_Try.
 DR Pfam: PF00089; trypsin; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRODOM: PD000395; Kringle; 1.
 DR SMART: SM00181; EGF; 1.

DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TYP_SPC; 1.

Query Match 9.5%; Score 132; DB 1; Length 431;
Best Local Similarity 32.7%; Pred. No. 0.00066;
Matches 33; Conservative 13; Mismatches 33; Indels 22; Gaps 5;

QY 24 CFMDNGHLYREDQSPAPGLRCLNWLDA-----OSGLASAPVSGAGNHSYCRNPDEDP 76
Db 70 CYEGNGHLYRKGKASTDTMGRRPCLPMSATVLOQTYHAHRSDALQLGLGKNYCRNPD-NR 128
QY 77 RGPWCYVSGEAGVPEKRP-----CEDLRCPETTSQAL 108
Db 129 RRPWCYV--QVGL--KPLVQECMVHDCADGKKSSPPEL 164

Search completed: April 7, 2003, 08:42:48
Job time : 13 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 08:39:12 ; Search time 33 Seconds

(Without alignments)
1635.889 Million cell updates/sec

Title: US-10-057-951-2_COPY_2_263

Sequence: 1 LLAWQAFIVSNMLAEAYG.....PVDPQSGSTPLMGAGTSGA 262

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_TODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_RVIRUS:*
16: SP_BACTERIAD:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------|---------------------|
| 1 | 1388 | 100.0 | 263 | 4 Q96FE7 | Q96FE7 homo sapien |
| 2 | 1385 | 99.8 | 263 | 4 O00318 | Q00318 homo sapien |
| 3 | 167 | 12.0 | 562 | 6 O8S023 | O8S023 sus scrofa |
| 4 | 154.5 | 11.1 | 516 | 4 Q9BU99 | Q9BU99 homo sapien |
| 5 | 150 | 10.8 | 653 | 11 Q8VCS4 | Q8VCS4 mus musculu |
| 6 | 147.5 | 10.6 | 559 | 11 Q91VP2 | Q91VP2 mus musculu |
| 7 | 142 | 10.2 | 810 | 4 Q1S146 | Q1S146 homo sapien |
| 8 | 139.5 | 10.1 | 385 | 5 Q25101 | Q25101 hermania m |
| 9 | 137 | 9.9 | 716 | 11 P70521 | P70521 rattus norv |
| 10 | 136.5 | 9.8 | 420 | 13 Q90504 | Q90504 eptactetus |
| 11 | 135 | 9.7 | 812 | 11 Q9R0W3 | Q9R0W3 rattus norv |
| 12 | 134.5 | 9.7 | 395 | 4 Q9BZM1 | Q9BZM1 homo sapien |
| 13 | 134.5 | 9.7 | 704 | 13 Q90865 | Q90865 gallus galli |
| 14 | 134 | 9.7 | 716 | 11 Q91XG8 | Q91XG8 mus musculu |
| 15 | 131.5 | 9.5 | 313 | 13 Q9PU78 | Q9PU78 crocodylus |
| 16 | 131 | 9.4 | 154 | 4 Q96S88 | Q96S88 homo sapien |

| | | | | | |
|----|-------|-----|------|-----------|---------------------|
| 17 | 131 | 9.4 | 608 | 13 Q9PTW7 | Q9PTW7 struthio ca |
| 18 | 130.5 | 9.4 | 717 | 13 P70006 | P70006 xenopus lae |
| 19 | 129 | 9.3 | 616 | 6 Q97507 | Q97507 sus scrofa |
| 20 | 128.5 | 9.3 | 806 | 6 Q18783 | Q18783 macropus eu |
| 21 | 127 | 9.1 | 157 | 6 Q9TVAS | Q9TVAS bos taurus |
| 22 | 126 | 9.1 | 716 | 13 Q91691 | Q91691 xenopus lae |
| 23 | 125.5 | 9.0 | 560 | 4 Q14520 | Q14520 homo sapien |
| 24 | 123 | 8.9 | 728 | 6 Q9BH09 | Q9BH09 felis silve |
| 25 | 121.5 | 8.8 | 399 | 4 Q96GL8 | Q96GL8 homo sapien |
| 26 | 121.5 | 8.8 | 420 | 4 Q9BTP9 | Q9BTP9 homo sapien |
| 27 | 119 | 8.6 | 334 | 6 Q46507 | Q46507 papio hamad |
| 28 | 119 | 8.6 | 710 | 13 Q91402 | Q91402 xenopus. lae |
| 29 | 117.5 | 8.5 | 812 | 11 Q91WJ5 | Q91WJ5 mus musculu |
| 30 | 114.5 | 8.2 | 113 | 4 Q9UR7 | Q9UR7 homo sapien |
| 31 | 114 | 8.2 | 113 | 4 Q9UR5 | Q9UR5 homo sapien |
| 32 | 114 | 8.2 | 202 | 13 Q90675 | Q90675 gallus galli |
| 33 | 114 | 8.2 | 452 | 13 Q90Y90 | Q90Y90 xenopus lae |
| 34 | 114 | 8.2 | 607 | 13 Q91001 | Q91001 gallus galli |
| 35 | 112.5 | 8.1 | 359 | 6 Q8WNR1 | Q8WNR1 canis famli |
| 36 | 112 | 8.1 | 567 | 4 Q13208 | Q13208 homo sapien |
| 37 | 112 | 8.1 | 594 | 5 P91823 | P91823 caenorhabdi |
| 38 | 112 | 8.1 | 1145 | 5 Q9BKL8 | Q9BKL8 aplysia cal |
| 39 | 111.5 | 8.0 | 597 | 11 Q35727 | Q35727 mus musculu |
| 40 | 111 | 8.0 | 726 | 13 Q90978 | Q90978 gallus galli |
| 41 | 110.5 | 8.0 | 685 | 5 Q24488 | Q24488 drosophila |
| 42 | 110.5 | 8.0 | 2358 | 16 Q911V8 | Q911V8 streptomyce |
| 43 | 108 | 7.8 | 215 | 13 Q42341 | Q42341 gallus galli |
| 44 | 106 | 7.6 | 378 | 13 Q90WP0 | Q90WP0 trachemys s |
| 45 | 105.5 | 7.6 | 709 | 13 Q90ZNE | Q90ZNE brachydantio |

ALIGNMENTS

RESULT 1

Q96FE7 PRELIMINARY; PRT; 263 AA.
AC Q96FE7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Unknown (protein for MGC:17330).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DDb databases.
DR EMBL, BC011049; AAH1049.1; -.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle_1.
DR ProDom; PD000395; Kringle_1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR PROSITE; PS00070; KRINGLE_2; 1.
SQ SEQUENCE 263 AA; 28234 MW; 197C3EE888FA242 CRC64;

Query Match 100.0%; Score 1388; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 2.6e-117;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-----|
| QY | 1 | LLAWQAFIVSNMLAEAYSGGCFWNGHLYREDQTSPPGIRCLINWIDAOGLASAPV | 60 |
| DB | 2 | LLAWQAFIVSNMLAEAYSGGCFWNGHLYREDQTSPPGIRCLINWIDAOGLASAPV | 61 |
| QY | 61 | SGAGNSYCRNDEDEPPRCPCYSGEAGVPEKPCEDLRCPETTSOALPAFTTEIOEASE | 120 |
| DB | 62 | SGAGNSYCRNDEDEPPRCPCYSGEAGVPEKPCEDLRCPETTSOALPAFTTEIOEASE | 121 |
| QY | 121 | GGGADEVQVFAPNALPARRSFAAVQPVYIGISQVRMNSKEKDLCTGVYVGIYMTYI | 180 |

Db 122 GPGADEVGFAPANALPARSEAAAVQPVIGISQVRVMSKEKKDGLTGLVYLITGMWYII 181
 QY 181 IAIAGGIIIGSYKRGDKLEQDHQKVCEREMQRTLPASFTNPTCEIIVDEKTVVHTS 240
 Db 182 IAIAGGIIIGSYKRGDKLEQDHQKVCEREMQRTLPASFTNPTCEIIVDEKTVVHTS 241
 QY 241 QTPVDPQEGSTPLMGQAGTPGA 262
 Db 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 2

000318 PRELIMINARY; PRT; 263 AA.
 AC 000318.
 DT 01-JUN-1997 (TREMBlrel. 04, Created)
 DT 01-JUN-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE WUGSC:DJ51SN1.2 protein.
 GN WUGSC:DJ51SN1.2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Du Z., Scheet P., Harper M.,
 RT "The sequence of H. sapiens PAC clone Rp3-51SN1."
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE FROM N.A.
 RA Waterston R.
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC002073; AAB54054.1; -.
 DR HSSP; P00749; 1KD0.
 DR InterPro; IPR000001; Kringle.
 DR Pfam; PF00051; Kringle_1.
 DR PRINTS; PR00018; KRINGLE; FALSE_NEG.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 SQ SEQUENCE 263 AA; 28248 MW; 197C3EEB854242 CRC64;

Query Match 99.8%; Score 1385; DB 4; Length 263;
 Best Local Similarity 99.6%; Pred. No. 4.9e-117;
 Matches 261; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIAWVQAFIVGNMLLAELVYSGGCFMNDGHLVREDQTSAPRGIRCLNMLDAQSGIAPV 60
 Db 2 LIAWVQAFIVGNMLLAELVYSGGCFMNDGHLVREDQTSAPRGIRCLNMLDAQSGIAPV 61
 QY 61 SGAGNHSYCRNPDEDPGRGFWCYSGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIQEASE 120
 Db 62 SGAGNHSYCRNPDEDPGRGFWCYSGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIQEASE 121
 QY 121 GPGADEVGFAPANALPARSEAAAVQPVIGISQVRVMSKEKKDGLTGLVYLITGMWYII 180
 Db 122 GPGADEVGFAPANALPARSEAAAVQPVIGISQVRVMSKEKKDGLTGLVYLITGMWYII 181
 QY 181 IAIAGGIIIGSYKRGDKLEQDHQKVCEREMQRTLPASFTNPTCEIIVDEKTVVHTS 240
 Db 182 IAIAGGIIIGSYKRGDKLEQDHQKVCEREMQRTLPASFTNPTCEIIVDEKTVVHTS 241
 QY 241 QTPVDPQEGSTPLMGQAGTPGA 262
 Db 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 3

08SQ23 PRELIMINARY; PRT; 562 AA.
 ID 08SQ23.
 AC 08SQ23;

DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE T-plasminogen activator.
 OS Sus scrofa (pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 OC NCBI_TaxID=9823;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=EMFAMEL ORGAN;
 RA Ding Y., Xue J., Bartlett J.D.;
 RT "T-plasminogen activator in tooth tissues."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF364605; AAM00297.1; -.
 SQ SEQUENCE 562 AA; 63668 MW; F9E6B4C77CB101E8 CRC64;

Query Match 12.0%; Score 167; DB 6; Length 562;
 Best Local Similarity 36.6%; Pred. No. 1.2e-06;
 Matches 41; Conservative 12; Mismatches 33; Indels 26; Gaps 6;

QY 24 CFPMNGHLVREDQTSAPRGIRCLNMLDAQSGIAPV-----GAGNHSYCRNPDE 74
 Db 127 CYEDGTYRTGWSTYSSGAEVNM--NTSGIASMPYNGRRPDAVKGLGHNHCRNPDK 184
 QY 75 DPRGFWCYV--SGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIQEASGPGAD 125
 Db 185 DSK-FWCYIFKAKKSPD-----FC-----STPACTKEKECYTGKGLD 222

RESULT 4

09B099 PRELIMINARY; PRT; 516 AA.
 AC 09B099.
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Similar to plasminogen activator, tissue.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=SKIN;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.

CC EMBL; BC002795; AA02795.1; -.
 DR HSSP; P00750; 1ASH.
 DR MEROPS; S01.232; -.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF; 1.
 DR Pfam; PF00051; kringle; 2.
 DR PRINTS; PR00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00001; EGF-like; 1.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_2.
 DR PROSITE; PS50070; KRINGLE_2; 2.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.


```

DE Plasmidogen precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Browne M.J., Chapman C.G., Dodd I., Carey J.E., Lawrence G.M.P.,
RT Mitchell D., Robinson J.H.,
RT "Expression of recombinant human plasminogen and aglycoplasminogen in
RT HeLa cells." 0:0-0(1991).
RL Fldrnolysls 0:0-0(1991).
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
DR EMBL; M74220; AAA36451.1; -.
DR HSRP; P00747; 2PK4.
DR MEROPS; S01.233; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 5.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS00070; KRINGLE_2; 5.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR Hydrolase; Serine protease; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 810 PLASMINOGEN.
SQ SEQUENCE 810 AA; 90555 MW; B05C7D4BD020B3C CRC64;

Query Match 10.2%; Score 142; DB 4; Length 810;
Best Local Similarity 37.2%; Pred. No. 0.00035;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 24 CFMDNGHLYREDQTSPPAGRLCLNWLDA-----QGLASAPVSGAGNHSYCRNPDEDPG 78
DB 103 CFTGNGKRYRGCTGSKTKYNGGITCKWMSSTSPHRPRFSPTATPSEGL-ENYCRNPDDNDPOG 161
QY 79 PWCYVSGEAGVPEKR--PCEDLRCE 102
DB 162 PWCYTTD---PEKRYDCILCECE 163

RESULT 8
Q25101 PRELIMINARY; PRT; 385 AA.
ID Q25101;
AC Q25101;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Serine proteinase.
GN HMRSP;
OS Herdmania momus.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Scolidobranchia; Eurythoe; Herdmania.
OX NCBI_TaxID=7733;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=CURVATA;
RA Arnold J.M., Kennett C., Lavín M.F.;
RT "Translational expression of a novel serine protease in the ectoderm of

```

```

RT the ascidian Herdmania momus during development.";
RL Dev. Gene Evol. 206:455-465(1997).
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
DR EMBL; U63177; AAB6650.1; -.
DR HSRP; P00763; IDPO.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR Hydrolase; Serine protease.
SQ SEQUENCE 385 AA; 42935 MW; BFB1D05D5232BE6A0 CRC64;

Query Match 10.1%; Score 139.5; DB 5; Length 385;
Best Local Similarity 20.2%; Pred. No. 0.00023;
Matches 70; Conservative 47; Mismatches 127; Indels 103; Gaps 14;

QY 1 LLAW--VOAFVSNMLAEAYSGGCF--WDNGHLYREDQTSPPAGRLCLNWLDAOS----- 53
DB 7 LVWITLNGFVSN-----SECFDENPESTQGLASITLGLGFCGSM-DLQTPHGH 56
QY 54 --GLASAPVSGAGNHSYCRNPDEDPGWCYVSGE-----AGVPEKRPCEDLR 99
DB 57 KYTSNTPVNSGLAGNNYCRNPDDMGKPCYTNTEFMKWDYCDIPICSNPVTLPSTIE 116
QY 100 CPETT---SQALPAFTTEIQEASGEG--GADEVQ----- 128
DB 117 CGKTEPLSDATKGYDLKQSKATNPVHVGDTVTHGSIPIWVSLRLKELRHFCGSI 176
QY 129 -----VFAPNALPASEAQAQPVIGISQVRMNSKEKDLGLTGLV----- 172
DB 177 LNRWILTAHICIKKQPKKYLALIGDYDRIGDYFSMK---VGRILFNHEKNPT 232
QY 173 --GITWVVIILAGIILGYSYKRGDLKEQHDQVC-----EREMQRITL 217
DB 233 FENDITLAKMDTISIALTIFGQSVPPANKVPAAKSKIVSGMDTKGTQDVKLQNVTL 292
QY 218 PLASATNPCTGIVDEKTV---VVHTSQTPVDPOGSGTPLMGQAGTP 260
DB 229 PWSF--KLCKKLYSKVGAAPVFTSLCAAYKKGKDCGSGSGGP 337

RESULT 9
P70521 PRELIMINARY; PRT; 716 AA.
ID P70521;
AC P70521;
DT 01-FEB-1997 (TREMblrel. 02, Created)
DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Macrophage stimulating protein precursor.
GN MSP;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RC MEDLINE=97011126; PubMed=8658136;
RA Ohsheiro K., Iwama A., Matsuno K., Ezaki T., Sakamoto O., Hamaguchi I.,
RA Takasu N., Suda T.;
RT "Molecular cloning of Rat Macrophage-stimulating protein and its

```

RT involvement in the Male Reproductive System.";
RL Biochem. Biophys. Res. Commun. 227:273-280(1996).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
DR EMBL; X95096; CAA64473.1; -.
DR HSSP; P00747; IKRN.
DR MEROPS; S01.975; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringles.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringles; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringles; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR PROSITE; PS00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR KMW; KMW000000; Serine protease; Signal.
FT SIGNAL 1 31 POTENTIAL.
SQ SEQUENCE 716 AA; 80733 MW; 06B7DF3EF56D921F CRC64;
Query Match 9.9%; Score 137; DB 11; Length 716;
Best Local Similarity 27.3%; Pred. No. 0.0084;
Matches 48; Conservative 12; Mismatches 56; Indels 60; Gaps 8;
QY 1 LLAWVAFVSNMLALAEVAGS-----GCFMNGHLYREDQTSAPAGLCLMW 48
DB 80 LLEWTV-----HSLAQHLHSSCDLPQKQYVTCIMDNGASRGIVATAGLGCGQM 134
QY 49 ---LDAQSLAPVSGAGNHSYCNEDPEDPGPCVYS----- 84
DB 135 SRFPNDHKYTPPKNGL-EENFCRNPDGDRPGMVCYTNRSVAFQSCGIGKSCRAVCVM 193
QY 85 -----GEAGVPEK-RPCE-DLRCPEP-----TSQALPAFTTEIQEASEGP 122
DB 194 CNGEDYRGVDTYSGRECECRMDLQHPHSHPEKPPDKALKNYCRANDASERP 249

RESULT 10
Q90504 PRELIMINARY; PRT; 420 AA.
ID Q90504;
AC Q90504;
DT 01-NOV-1996 (TREMblrel. 01. Created)
DT 01-NOV-1996 (TREMblrel. 01. Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20. Last annotation update)
DE Thrombin.
OS Eptacretus stoutii (Pacific hagfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;
OC Myxiniidae; Eptacretinae; Eptacretus.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=92212913; PubMed=155783;
RA Banfield D.K., Macgillivray R.T.;
RT "Partial characterization of vertebrate prothrombin cDNAs:
RT amplification and sequence analysis of the B chain of thrombin from
RT nine different species.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=94223694; PubMed=7513365;
RA Banfield D.K., Irwin D.M., Walz D.A., Macgillivray R.T.;
RT "Evolution of prothrombin: isolation and characterization of the cDNAs
RT encoding chicken and hagfish prothrombin.";

RL J. Mol. Evol. 38:177-187(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Banfield D.K.;
RT Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
DR EMBL; M81393; AAA21620.1; -.
DR HSSP; P00734; IUVS.
DR MEROPS; S01.217; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringles.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringles; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringles; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR KMW; KMW000000; Serine protease.
SQ SEQUENCE 420 AA; 47888 MW; 64522AA21A57B67A CRC64;
Query Match 9.8%; Score 136.5; DB 13; Length 420;
Best Local Similarity 26.5%; Pred. No. 0.0047;
Matches 39; Conservative 16; Mismatches 77; Indels 15; Gaps 4;
QY 24 CFMNGHLYREDQTSAPAGLCLMWLDAQSLAPVSGAG-NHSYCRNPDGPGPCVY 82
DB 17 CVRRGRDVRGDNLTWTGKPCLPWRGYSNFPSPQTTNGLTSTNCRNPDGDSQWCY 76
QY 83 VSGEAGVP-----EKRPCEDLRCPEPETSQALPAFTTEIQEASEGPADEVQVAPANALPA 138
DB 77 TKGVSEGTVDVYQCLNYCE-----SGDIFEVGTDEVOYLSGRSGAAEKTLFPNPKTRGN 129
QY 139 RSRRAAVOPVIGISQVRVMSKEKKDL 165
DB 130 GEBEGCKRPFELQK--NDRSEDEL 153

RESULT 11
Q90503 PRELIMINARY; PRT; 812 AA.
ID Q90503;
AC Q90503;
DT 01-MAY-2000 (TREMblrel. 13. Created)
DT 01-MAY-2000 (TREMblrel. 13. Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20. Last annotation update)
DE Plasminogen protein precursor (EC 3.4.21.7).
GN PLASMINOGEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=91250378; PubMed=1645711;
RA Kanalas J.J., Makker S.P.;
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT receptor site for plasminogen.";
RL J. Biol. Chem. 266:10825-10829(1991).

```

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: AJ242649; CAB64014.1; -.
DR HSSP: P00747; 1PMK.
DR MEROPS: S01.233; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Ser_protease_Try.
DR InterPro: IPR001400; Somatostatin.
DR Pfam: PF00051; Kringle; 5.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 5.
DR SMART: SM00130; KR; 4.
DR SMART: SM00473; PAN; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS00021; KRINGLE_1; 5.
DR PROSITE: PS50070; KRINGLE_2; 5.
DR PROSITE: PS00338; SOMATOSTATIN_2; UNKNOWN_1.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Serine protease; Signal.
FT SIGNAL 1 19
FT CHAIN 20 812 PLASMINOGEN.
SQ SEQUENCE 812 AA; 90535 MW; 8C703C51410BCE9E CRC64;

Query Match 9.7%; Score 135; DB 11; Length 812;
Best Local Similarity 30.9%; Pred. No. 0.0015;
Matches 34; Conservative 14; Mismatches 48; Indels 14; Gaps 5;

QY 24 CFWDNGHLYREDQTSPPAGRLCLNWLDA-----QSGLASAPVSGAGNHSYCRNPDEPRG 78
DB 376 CYOQNGKSRGTSSTTGKCKQSGVSMTPSHSKTPANPPDGL-ENMYCNPNDORG 434
QY 79 PNCYVSGEAGVPERK--PCEDLRCPERTSQALPAFTTEIOEASBEGGADE 126
DB 435 PNCFTTD---PSVMEYCNKRCSEFTGGV--AESAIIVQVSAFGTSE 478

RESULT 12
Q9BZM1 PRELIMINARY; PRT; 395 AA.
ID Q9BZM1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Neocatal thymolytic agent alpha-form (fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dou D.;
RT "A brain-type plasminogen activator.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: AF260825; AAK11956.1; -.
DR HSSP: P00750; 1PK2.
DR MEROPS: S01.232; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000083; Fibrinolytic.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00039; PAN; 1.
DR Pfam: PF00051; Kringle; 1.
DR Pfam: PF00089; trypsin; 1.

```

```

DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 1.
DR SMART: SM00058; FN1; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS01253; FIBRINOLYTIN_1; 1.
DR PROSITE: PS00021; KRINGLE_1; UNKNOWN_1.
DR PROSITE: PS50070; KRINGLE_2; 1.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Serine protease.
FT NOW TER 395
SQ SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match 9.7%; Score 134.5; DB 4; Length 395;
Best Local Similarity 32.3%; Pred. No. 0.00066;
Matches 32; Conservative 13; Mismatches 33; Indels 21; Gaps 4;

QY 3 AMVOAFIVSNML-----LAEAVSGGCFWDNGHLYREDQTSPPAGRLCLNWLDA-----48
DB 18 SMLPVLRSNRVEYCMNGSGRACSEGNDCYFGSGAYRGTHLTSGASCLPMNSMIL 77
QY 49 ---LDQSGLASAPVSGAGNHSYCRNPDEPRGWCY 83
DB 78 IGKYYTAQN--PSAQLGKGNKYNCRNPDDGAK-PWCHV 113

RESULT 13
Q90865 PRELIMINARY; PRT; 704 AA.
ID Q90865;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hepatocyte growth factor-like/macrophage stimulating protein.
GN HGFL/MSP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RN TISSUE=Liver;
RX MEDLINE=96029010; PubMed=7554499;
RA Thery C., Sharpe M.J., Bailey S.J., Stern C.D., Cheraudi E.;
RT "Expression of HGF/SF, HGFL/MSP and c-met suggests new functions
RT during early chick development.";
RT Dev. Genet. 17:90-101(1995)
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: X84043; CAAS862.1; -.
DR HSSP: P00747; 1CEA.
DR MEROPS: S01.977; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00051; Kringle; 4.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 4.
DR SMART: SM00130; KR; 4.
DR SMART: SM00473; PAN; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS00021; KRINGLE_1; 4.
DR PROSITE: PS50070; KRINGLE_2; 4.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.

```

KM Hydrolyase; Serine protease.
SQ SEQUENCE 704 AA; 79341 MW; CAB08DCA1367C37 CRC64;

Query Match 9.7%; Score 134.5; DB 13; Length 704;
Best Local Similarity 28.8%; Pred. No. 0.0014;

Matches 32; Conservative 11; Mismatches 45; Indels 23; Gaps 4;

QY 24 CFWMNGHYREDQTSPPAPGLRCINWLDASGLASAP-----VSGAGNHSYCRNPED 75
DB 108 CIYANGTSYGRDTPRGLRCQHW-----QATPHDHRFLPSLRNGLIENYCRNPED 161
QY 76 PRGWCYV-----SGAGVPEKRPCEDLRCPETTSQALPATTTIOGASE 120
DB 162 KRGWCYTVDPNVVHOSGCI--KCEDAVCMTCNGEDYRGVHDHTESGTE 209

RESULT 14

Q91XG8 PRELIMINARY; PRT; 716 AA.
AC Q91XG8; 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE Hepatocyte growth factor-like.
GN HGFL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC010551; AA010551.1; -
DR MGI; MGI:96080; Hgfl.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; Kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRODOM; PD000395; Kringle; 4.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_4.
DR PROSITE; PS50070; KRINGLE_2; 4.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
KW Hydrolyase; Serine protease.
SQ SEQUENCE 716 AA; 80693 MW; 12474C48A7D4B46D CRC64;

Query Match 9.7%; Score 134; DB 11; Length 716;
Best Local Similarity 28.2%; Pred. No. 0.0016;
Matches 42; Conservative 9; Mismatches 46; Indels 52; Gaps 7;

QY 1 LLAWQAFLVSNMLAAYSG-----GCFMNGHYREDQTSPPAPGLRCINW 48
DB 80 LLPWTQ-----HSHTQLYHSLSLCHLFQKKDYRTICINDNGVSYRGVARTAGLPQAW 134
QY 49 ---LDAOSGLASAVSAGNHSYCRNPDEDPGWCYV----- 84
DB 135 SRPFPNKHKTPTPKNGL-BENFCRNPDDGPRGWCYTTNRSVRQSGIKTCREAVCVL 193
QY 85 -----GEAGVPEK-RPCE--DLRCPET 103
DB 194 CNGEDYRGVVDVTESGRCQRMWLDQHPHS 222

RESULT 15

Q9PU78 PRELIMINARY; PRT; 313 AA.
AC Q9PU78; 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE Hepatocyte growth factor-like protein (Fragment).
OS Crocodylus niloticus (Nile crocodile) (African crocodile).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Crocodylinae; Crocodylus.
RX NCBI_TaxID=6501;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA MEDLINE=20022983; PubMed=1055283;
RA Hughes S., Zelus D., Mouchiroud D.;
RT "Warm-blooded isochore structure" in Nile crocodile and turtle.";
RT Mol. Biol. Evol. 16:1521-1527 (1999).
DR EMBL; AJ011396; CAB56422.1; -
DR HSSP; P00747; IHPV.
DR MEROPS; S01.977; -
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; Kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00019; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
KW Hydrolyase; Serine protease.
FT NON_TER 1 1
FT NON_TER 313 313
SQ SEQUENCE 313 AA; 34793 MW; 8E084704958B5AA2 CRC64;

Query Match 9.5%; Score 131.5; DB 13; Length 313;
Best Local Similarity 29.8%; Pred. No. 0.00092;
Matches 28; Conservative 11; Mismatches 42; Indels 13; Gaps 4;

QY 24 CFWMNGHYREDQTSPPAPGLRCINWLDASGLAS-APVSGAGNH---SYCRNPDDPRGP 79
DB 30 CYHGNGELYRGHTSKTRKGVTCQMSQSPHYPOISPTTHPAHLIDENYCRNPDDSHGP 89
QY 80 WCYVSGE-----AGVPEKRPCEDLRCPETTSQA 107
DB 90 WCYTMDPRTPPDYCGI--KPCAGDKIPVLENA 120

Search completed: April 7, 2003, 08:43:28
Job time : 35 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 08:34:02 ; Search time 37 Seconds

(without alignments)
943.558 Million cell updates/sec

Title: US-10-057-951-2_COPY_2_263

Perfect score: 1388
Sequence: 1 LIAWVQAFVISMMLAEAYG.....PVDPEGSTPLMGQAGTTPGA 262

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/emb1/AA2000.DAT.*
21: /SIDS2/gcgdata/geneseq/emb1/AA2001.DAT.*
22: /SIDS2/gcgdata/geneseq/emb1/AA2002.DAT.*
23: /SIDS2/gcgdata/geneseq/emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------|
| 1 | 1388 | 100.0 | 263 | 20 | AAV05219 |
| 2 | 1388 | 100.0 | 263 | 20 | AAH87769 |
| 3 | 1388 | 100.0 | 263 | 22 | AAH00100 |
| 4 | 1385 | 99.4 | 263 | 23 | AAH86149 |
| 5 | 1379 | 99.4 | 263 | 21 | AAH43237 |
| 6 | 1374 | 99.0 | 263 | 22 | AAH93748 |
| 7 | 1162.5 | 83.8 | 286 | 20 | AAV05220 |
| 8 | 322 | 23.2 | 66 | 22 | ABH37905 |
| 9 | 322 | 23.2 | 66 | 22 | ABH23159 |
| 10 | 322 | 23.2 | 66 | 22 | AAH58537 |

| | | | | | | |
|----|-------|------|-----|----|----------|--------------------|
| 11 | 322 | 23.2 | 66 | 22 | AAH71037 | Human bone marrow |
| 12 | 322 | 23.2 | 66 | 22 | AAH18800 | Peptide #5314 enco |
| 13 | 322 | 23.2 | 66 | 23 | AAH31314 | Peptide #5315 enco |
| 14 | 322 | 23.2 | 66 | 23 | ABG40828 | Human peptide enco |
| 15 | 301 | 21.7 | 56 | 20 | AAH12615 | Human 5' EST gene |
| 16 | 300 | 21.6 | 55 | 20 | AAH12397 | Human 5' EST gene |
| 17 | 196 | 14.1 | 39 | 19 | AAH72641 | Nervous glia cell |
| 18 | 192 | 13.8 | 39 | 19 | AAH72640 | Nervous glia cell |
| 19 | 160.5 | 11.6 | 527 | 13 | AAH20217 | t-PA analogue exp |
| 20 | 160.5 | 11.6 | 527 | 13 | AAH20220 | t-PA analogue exp |
| 21 | 160.5 | 11.6 | 527 | 13 | AAH20218 | t-PA analogue exp |
| 22 | 160.5 | 11.6 | 527 | 13 | AAH20219 | t-PA analogue exp |
| 23 | 160.5 | 11.6 | 527 | 13 | AAH20219 | t-PA analogue exp |
| 24 | 160.5 | 11.6 | 527 | 13 | AAH20219 | t-PA analogue exp |
| 25 | 160.5 | 11.6 | 527 | 13 | AAH20219 | t-PA analogue exp |
| 26 | 160.5 | 11.6 | 527 | 13 | AAH20219 | t-PA analogue exp |
| 27 | 159.5 | 11.5 | 527 | 13 | AAH20222 | t-PA analogue exp |
| 28 | 159.5 | 11.5 | 527 | 13 | AAH20222 | t-PA analogue exp |
| 29 | 159.5 | 11.5 | 527 | 13 | AAH20221 | t-PA analogue exp |
| 30 | 157.5 | 11.3 | 527 | 19 | AAH54154 | t-PA mutant (N1425 |
| 31 | 157.5 | 11.3 | 527 | 19 | AAH54157 | t-PA mutant (N1425 |
| 32 | 157.5 | 11.3 | 562 | 10 | AAH94379 | Amino acid sequenc |
| 33 | 157.5 | 11.3 | 562 | 10 | AAH94380 | Amino acid sequenc |
| 34 | 156.5 | 11.3 | 439 | 16 | AAH68851 | Delta 2-89 tissue |
| 35 | 156.5 | 11.3 | 483 | 16 | AAH70889 | Human tissue PA va |
| 36 | 156.5 | 11.3 | 483 | 16 | AAH70877 | Human tissue PA va |
| 37 | 156.5 | 11.3 | 483 | 16 | AAH70878 | Human tissue PA va |
| 38 | 156.5 | 11.3 | 483 | 16 | AAH70879 | Human tissue PA va |
| 39 | 156.5 | 11.3 | 483 | 16 | AAH70880 | Human tissue PA va |
| 40 | 156.5 | 11.3 | 483 | 16 | AAH70881 | Human tissue PA va |
| 41 | 156.5 | 11.3 | 483 | 16 | AAH70882 | Human tissue PA va |
| 42 | 156.5 | 11.3 | 483 | 16 | AAH70883 | Human tissue PA va |
| 43 | 156.5 | 11.3 | 483 | 16 | AAH70884 | Human tissue PA va |
| 44 | 156.5 | 11.3 | 483 | 16 | AAH70885 | Human tissue PA va |
| 45 | 156.5 | 11.3 | 483 | 16 | AAH70886 | Human tissue PA va |

ALIGNMENTS

| | | |
|----------|----------------------------------|---|
| RESULT 1 | AAV05219 | standard; Protein, 263 AA. |
| ID | AAV05219 | standard; Protein, 263 AA. |
| XX | AAV05219 | |
| AC | AAV05219 | |
| XX | 17-JUN-1999 | (first entry) |
| DT | XX | |
| DE | XX | Kringle1 protein sequence. |
| XX | XX | |
| KM | XX | Kringle1; cancer; inflammation; autoimmunity; rheumatoid arthritis; AIDS; |
| KM | XX | CNS inflammation; cerebellar degeneration; Alzheimer's disease; asthma; |
| KM | XX | Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis; |
| KM | XX | neurological abnormality; ischaemia reperfusion injury; ischemic injury; |
| KM | XX | cardiovascular disease; kidney disease; liver disease; aplastic anemia; |
| KM | XX | myocardial infarction; hypotension; hypertension; allergy; infection; |
| KM | XX | myelodysplastic syndrome; haematologic abnormality; diagnosis; therapy; |
| KM | XX | male pattern baldness. |
| OS | XX | Homo sapiens. |
| XX | XX | |
| PN | MO9911786-A1. | |
| XX | XX | |
| PD | 11-MAR-1999. | |
| XX | XX | |
| PF | 02-SEP-1998; | 98MO-US18270. |
| XX | XX | |
| PR | 01-SEP-1998; | 98US-0144889. |
| XX | XX | |
| PR | 02-SEP-1997; | 97US-0056037. |
| XX | XX | |
| PA | (SMIK) SMITHKLINE BEECHAM CORP. | |
| XX | XX | |
| PI | Albone EF, Kikly KK; | |

```

XX WPI: 1999-214707/18.
DR N-PSDB; MAX28354.
XX
PT New kringle1 polypeptides and polynucleotides
XX
PS Claim 1; Page 31-32; 42pp; English.
XX
CC This sequence is a Kringle1 polypeptide of the invention.
CC The Kringle1 polypeptides (I) are used to screen for agonists and
CC antagonists. Agonists are used to treat subjects in need of enhanced
CC activity or expression of (I). Antagonists are used to treat subjects
CC having need to inhibit the activity or expression of (I). The methods can
CC be used to treat conditions such as cancer, inflammation, autoimmunity,
CC allergy, asthma, rheumatoid arthritis, CNS inflammation, cerebellar
CC degeneration, Alzheimer's disease, Parkinson's disease, multiple
CC sclerosis, amyotrophic lateral sclerosis, head injury damage and other
CC neurological abnormalities, ischaemia reperfusion injury, cardiovascular
CC disease, kidney disease, liver disease, ischaemic injury, myocardial
CC infarction, hypotension, hypertension, AIDS, myelodysplastic syndromes
CC and other haematologic abnormalities, aplastic anaemia, male pattern
CC baldness, and bacterial, fungal, protozoan and viral infections. The
CC Kringle1 polypeptides may also be used to generate antibodies.
CC Determining the presence or absence of mutations in, and analysing for
CC the presence or absence of expression of, Kringle1 polynucleotides can be
CC used to diagnose a disease or susceptibility to a disease related to
CC expression or activity of Kringle1 proteins. The polynucleotides may also
CC be used for chromosome identification, and mapping.
XX
SQ Sequence 263 AA;
XX
Query Match 100.0%; Score 1388; DB 20; Length 263;
Best Local Similarity 100.0%; Pred. No. 2, 1e-13;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 LLAHQATLVSNMLLAAYSGGCTGWDGHIYREDQTSPPAGRCINMLDQSGLASAPV 60
DB 2 LLAHQATLVSNMLLAAYSGGCTGWDGHIYREDQTSPPAGRCINMLDQSGLASAPV 61
QY 61 SGAGHSTYCRNPDDPRGMCVSGAGVPEKPCEDLRCPETTSQALPAFTTIOQAS 120
DB 62 SGAGHSTYCRNPDDPRGMCVSGAGVPEKPCEDLRCPETTSQALPAFTTIOQAS 121
QY 121 GFGADEVQVFPANALPARSEAAAVQPIGISOVRNNSKEKKDGLGYVLGITMVA 180
DB 122 GFGADEVQVFPANALPARSEAAAVQPIGISOVRNNSKEKKDGLGYVLGITMVA 181
QY 181 IAIAGIILGYSGKQKDLKEQDKVCEMORITPLSAFTNPCEIVDEKTVVHTS 240
DB 182 IAIAGIILGYSGKQKDLKEQDKVCEMORITPLSAFTNPCEIVDEKTVVHTS 241
QY 241 QTPVDPQEGSTPLMGQAGTPGA 262
DB 242 QTPVDPQEGSTPLMGQAGTPGA 263
XX
RESULT 2
AAW87769
ID AAW87769 standard; Protein; 263 AA.
XX
AC AAW87769;
XX
DT 29-MAR-1999 (first entry)
XX
DE Human tissue plasminogen activator-like protease t-PAAP.
XX
KM Tissue plasminogen activator-like protease; t-PAAP; human;
KM circulatory system-related disorder; blood clotting; stroke;
KM thrombosis; peripheral arterial occlusion; pulmonary embolism;
KM myocardiocirculation; diagnosis; therapy.
XX
OS Homo sapiens.
XX

```

```

FH Key Location/Qualifiers
FT Peptide 1..21
FT /label= Sig_peptide
FT 22..263
FT /label= Mac_protein
FT 25..84
FT /note= "kringle domain"
FT 85..263
FT /note= "protease domain"
FT 22..31
FT /note= "epitope-bearing region"
FT 35..44
FT /note= "epitope-bearing region"
FT 71..81
FT /note= "epitope-bearing region"
FT 91..107
FT /note= "epitope-bearing region"
FT 119..128
FT /note= "epitope-bearing region"
FT 138..147
FT /note= "epitope-bearing region"
FT 155..167
FT /note= "epitope-bearing region"
FT 193..203
FT /note= "epitope-bearing region"
FT 206..215
FT /note= "epitope-bearing region"
FT 227..237
FT /note= "epitope-bearing region"
FT 243..252
FT /note= "epitope-bearing region"
FH
FH W09854199-A1.
FH 03-DEC-1998.
FH
FH 27-MAY-1998; 98WO-US10728.
FH
FH 28-MAY-1997; 97US-0048000.
FH
FH (HUMA-) HUMAN GENOME SCI INC.
FH
FH Edner R, Moore PA, Ruben SM;
FH
FH WPI: 1999-070207/06.
FH N-PSDB; AAV99636.
FH
FH Claim 1; Page 56-57; 76pp; English.
FH
FH This is the amino acid sequence of tissue plasminogen activator-like
FH protease (t-PAAP), a novel member of the serine protease family
FH that shares sequence homology to human tissue plasminogen activator
FH (see AAW87769). The t-PAAP sequence was deduced from a cDNA clone
FH (see AAW99636) derived from activated monocytes. The 2.5 kb t-PAAP
FH message has also been detected in heart, brain, lung, placenta,
FH liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate,
FH testis, ovary, small intestine, colon and peripheral blood
FH leukocytes. Isolated nucleic acids encoding amino acids -21 to
FH 242, -20 to 242, 1-242, 4-63 (kringle domain) and 64-242 (protease
FH domain) of t-PAAP, or encoding epitope-bearing portions of t-PAAP,
FH are also claimed, as are recombinant vectors, host cells, and
FH methods for producing t-PAAP polypeptides. t-PAAP may be used to
FH detect and treat disorders related to the circulatory system, and
FH to identify agonists and antagonists of t-PAAP activity. The
FH homology between t-PAAP and tPA indicates that t-PAAP may be
FH involved in the regulation of normal and abnormal clotting
FH in e.g. stroke, deep-vein thrombosis, peripheral arterial
FH occlusion, pulmonary embolism and myocardiocirculation.
FH
FH Sequence 263 AA;

```



```

Query Match      100.0%; Score 1388; DB 20; Length 263;
Best Local Similarity 100.0%; Pred. No. 2.1e-119;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAIVQAFVLSNMMLLAERAYSGGCFMNDGHLVREDQTSPPAGRLCLNMLDNOGSLASAPV 60
DB 2 LLAIVQAFVLSNMMLLAERAYSGGCFMNDGHLVREDQTSPPAGRLCLNMLDNOGSLASAPV 61
QY 61 SGAGNHSYCRNPDEDPNGWCYSGEAGVPRKPCEDLRCPETTSQALPAFTTEIQEASE 120
DB 62 SGAGNHSYCRNPDEDPNGWCYSGEAGVPRKPCEDLRCPETTSQALPAFTTEIQEASE 121
QY 121 GPGADEVQVFPANALPFRSEAAVQPIGTSORVRNMSKKKDLGTLGYVLGITMMVII 180
DB 122 GPGADEVQVFPANALPFRSEAAVQPIGTSORVRNMSKKKDLGTLGYVLGITMMVII 181
QY 181 IAIAGGIIIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPTEIYDEKTVVHTS 240
DB 182 IAIAGGIIIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPTEIYDEKTVVHTS 241
QY 241 QTPVDPQEGSTPLMGQAGTPGA 262
DB 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 3
AAE00300
ID AAE00300 standard; Protein; 263 AA.
XX
AC AAE00300;
XX
DT 13-JUN-2001 (first entry)
DE Human tissue-plasminogen activator-like protease (t-PALP).
XX
XX Human; tissue-plasminogen activator-like protease; t-PALP;
XX therapy; vascular disease; stroke; deep vein thrombosis; keloid; asthma;
XX arterial occlusion; blood coagulation disorder; cerebroprotective;
XX autoimmune system disorder; human immunodeficiency syndrome; cytostatic;
XX rheumatoid arthritis; graft-versus-host disease; thyroiditis; cartilag;
XX insulin dependent diabetes; inflammatory eye disease; allergy; antiviral;
XX cardiovascular disease; heart disease; arrhythmia; myocardial ischemia;
XX hyperproliferative disorder; hypertrophic scar; neurological disease;
XX Creutzfeldt-Jakob syndrome; neurodegenerative disorder; antiinflammatory;
XX Alzheimer's disease; Parkinson's disease; immunosuppressive; antifungal;
XX infectious disease; drug screening; gene therapy; neuroprotective;
XX cancer; ophthalmological; antibacterial; vulnerary.
XX
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX Binding-site 1..165
XX /note= "Binds to FLAG polypeptide to form
XX t-PALP-FLAG fusion protein"
XX Peptide 1..21
XX /label= Signal_peptide
XX Domain 4..63
XX /label= Kringle_domain
XX Region 12..21
XX /note= "Conserved region"
XX Protein 22..263
XX /note= "Human mature tissue-plasminogen activator-like
XX protease (t-PALP); Binds to FLAG polypeptide to form
XX t-PALP-FLAG fusion protein"
XX Region 22..38
XX /note= "Conserved region"
XX Region 22..31
XX /note= "Epitope-bearing portion"
XX Region 35..44
XX /note= "Epitope-bearing portion"
XX Region 39..49
XX /note= "Conserved region"

```

```

FT Region 50..62
FT /note= "Conserved region"
FT Region 63..84
FT /note= "Conserved region"
FT Domain 64..242
FT /label= Protease_domain
FT Region 71..81
FT /note= "Epitope-bearing portion"
FT Region 85..97
FT /note= "Conserved region"
FT Region 91..107
FT /note= "Epitope-bearing portion"
FT Region 100..118
FT /note= "Conserved region"
FT Region 119..128
FT /note= "Epitope-bearing portion"
FT Region 119..127
FT /note= "Conserved region"
FT Region 128..143
FT /note= "Conserved region"
FT Region 138..147
FT /note= "Epitope-bearing portion"
FT Region 146..163
FT /note= "Conserved region"
FT Region 155..167
FT /note= "Epitope-bearing portion"
FT Region 164..180
FT /note= "Conserved region"
FT Region 186..200
FT /note= "Conserved region"
FT Region 193..203
FT /note= "Epitope-bearing portion"
FT Region 201..220
FT /note= "Conserved region"
FT Region 206..215
FT /note= "Epitope-bearing portion"
FT Region 221..236
FT /note= "Conserved region"
FT Region 227..237
FT /note= "Epitope-bearing portion"
FT Region 237..248
FT /note= "Conserved region"
FT Region 243..252
FT /note= "Epitope-bearing portion"
FT Region 249..263
FT /note= "Conserved region"
XX
XX MO200125252-A1.
XX
XX 12-APR-2001.
XX
XX 03-OCT-2000; 2000MO-US27239.
XX
XX 04-OCT-1999; 99US-0411977.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Moore PA, Ruben SM, Ebner R;
XX WPI: 2001-235402/24.
XX N-PSDB; AAD03460.
XX
XX New (gene encoding and antibody immunospecific for a)
XX tissue-plasminogen activator-like protease, useful for the diagnosis
XX and treatment of (cardio)vascular diseases, hyperproliferative
XX disorders, immune system disorders and cancers -
XX
XX Claim 17; Fig 1; 323pp; English.
XX
XX The present amino acid sequence is HMS1B42 clone human
XX tissue-plasminogen activator-like protease (t-PALP). The t-PALP
XX sequence and their (ant)agonists are useful for the diagnosis and
XX treatment of vascular diseases e.g. stroke, deep vein thrombosis and

```

arterial occlusion, blood coagulation disorders, (auto)immune system disorders e.g. human immunodeficiency syndrome, rheumatoid arthritis, graft-versus-host disease, thyroiditis, insulin dependent diabetes and inflammatory eye disease, allergic reactions e.g. asthma, cardiovascular diseases e.g. heart disease, arrhythmia and myocardial ischaemia, hyperproliferative disorders, cancers, hypertrophic scars and keloids, neurological diseases e.g. Creutzfeldt-Jakob syndrome, neurodegenerative disorders e.g. Alzheimer's disease and Parkinson's disease and infectious disease e.g. viral, bacterial and fungal infections. The t-PALP sequences are also useful for drug screening. The t-PALP nucleotides are useful as chromosome markers and are involved in gene therapy.

Sequence 263 AA;

Query Match 100.0%; Score 1388; DB 22; Length 263;
Best Local Similarity 100.0%; Pred. No. 2, 1e-119;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAWVQAFVLSNMMLAEAYVGGGCFWMDNGHLYREDQTSAPAGLCLNWLDAQSLASAPV 60
DB 2 LLAWVQAFVLSNMMLAEAYVGGGCFWMDNGHLYREDQTSAPAGLCLNWLDAQSLASAPV 61
QY 61 SGAGNHSTCRNPDEDPKPCVSGEAGVPEKRPCEDLRCEITTSQALPAFTTEIQEASE 120
DB 62 SGAGNHSTCRNPDEDPKPCVSGEAGVPEKRPCEDLRCEITTSQALPAFTTEIQEASE 121
QY 121 GPGADEVQVFPANALPARSEAAVQPVIGISQVRNMSKEKDLGTLGYVLGITMMVII 180
DB 122 GPGADEVQVFPANALPARSEAAVQPVIGISQVRNMSKEKDLGTLGYVLGITMMVII 181
QY 181 IAIAGGIIIGSYRKGDKLKEQHDQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTS 240
DB 182 IAIAGGIIIGSYRKGDKLKEQHDQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTS 241
QY 241 QTPVDPQEGSTPLMGQAGTPGA 262
DB 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 4

AAU86149 ID AAU86149 standard; Protein; 263 AA.

XX AAU86149;

DT 15-JUL-2002 (first entry)

DE Human PRO264 polypeptide.

KW Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;

KW Leukemia; neuronal disorder; stromal disorder; blastocoele disorder;

KW Cytostatic; neuroprotective.

OS Homo sapiens.

XX WO200153486-A1.

PD 26-JUL-2001.

PF 11-FEB-2000; 2000WO-US03565.

PR 08-MAR-1999; 99WO-US05028.

PR 11-MAR-1999; 99US-123972P.

PR 11-MAY-1999; 99US-133459P.

PR 02-JUN-1999; 99WO-US12252.

PR 22-JUN-1999; 99US-140650P.

PR 22-JUN-1999; 99US-140653P.

PR 20-JUL-1999; 99US-144758P.

PR 26-JUL-1999; 99US-145698P.

PR 28-JUL-1999; 99US-146222P.

PR 17-AUG-1999; 99US-149395P.

PR 31-AUG-1999; 99US-151689P.

PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28301.
PR 05-JAN-2000; 2000WO-US08219.

XX (GENTH) GENENTECH INC.

PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;

PI Masters SA, Pan J, Pitti RM, Roy MA, Smith V, Stone DM;

PI Matanabe CK, Wood WI;

DR WPI; 2002-205567/26.

XX N-PSDB; ABR40275.

PT Thirty five nucleic acids encoding PRO polypeptides, useful for

PT treating benign or malignant tumours, leukemias and lymphoid

PT malignancies, inflammatory, angiogenic and immunologic disorders -

PS Claim 61; Fig 44; 302pp; English.

XX The present invention relates to the isolation of novel human PRO
XX polypeptides and the polynucleotide sequences encoding them. The
XX PRO polypeptides, agonists, antagonists or anti-PRO antibodies are
XX useful for treating benign or malignant tumours (e.g. renal, kidney,
XX bladder, breast, etc), leukemias and lymphoid malignancies, other
XX disorders such as neuronal, glial, astrocytic, hypothalamic, glandular,
XX macrophagal, stromal and blastocoele disorders, inflammatory, immune
XX and angiogenic disorders. The polynucleotide sequences are also
XX useful in gene therapy. AAU86128-AAU86162 represent the human PRO
XX polypeptides of the invention.

SEQ Sequence 263 AA;

Query Match 99.8%; Score 1385; DB 23; Length 263;
Best Local Similarity 99.6%; Pred. No. 4e-119;
Matches 261; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAWVQAFVLSNMMLAEAYVGGGCFWMDNGHLYREDQTSAPAGLCLNWLDAQSLASAPV 60
DB 2 LLAWVQAFVLSNMMLAEAYVGGGCFWMDNGHLYREDQTSAPAGLCLNWLDAQSLASAPV 61
QY 61 SGAGNHSTCRNPDEDPKPCVSGEAGVPEKRPCEDLRCEITTSQALPAFTTEIQEASE 120
DB 62 SGAGNHSTCRNPDEDPKPCVSGEAGVPEKRPCEDLRCEITTSQALPAFTTEIQEASE 121
QY 121 GPGADEVQVFPANALPARSEAAVQPVIGISQVRNMSKEKDLGTLGYVLGITMMVII 180
DB 122 GPGADEVQVFPANALPARSEAAVQPVIGISQVRNMSKEKDLGTLGYVLGITMMVII 181
QY 181 IAIAGGIIIGSYRKGDKLKEQHDQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTS 240
DB 182 IAIAGGIIIGSYRKGDKLKEQHDQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTS 241
QY 241 QTPVDPQEGSTPLMGQAGTPGA 262
DB 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 5

AAAB43237 ID AAB43237 standard; Protein; 263 AA.

XX AAB43237;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF3001 polypeptide sequence SEQ ID NO:6002.

KW Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;

KW vulnerable; antiproliferative; antiparkinsonian; neurotrophic; neuroprotective;

KW anticonvulsant; osteopathic; antiarrhythmic; immunosuppressant; cardiac;

KM immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KM hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KM antiviral; antibacterial; antifungal; antirheumatic; antichryd;
 KM antiaiemic; gene therapy; cancer; proliferative disorder; hyperetation;
 KM neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KM cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KM cholesterol ester storage; systemic lupus erythematosus; infection;
 KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KM allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KM bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KM thrombosis; contraceptive.
 OS Homo sapiens.
 XX W0200058473-A2.
 XX PD 05-OCT-2000.
 XX PF 31-MAR-2000; 2000WO-US08621.
 XX PR 31-MAR-1999; 99US-0127607.
 XX PR 02-APR-1999; 99US-0127636.
 XX PR 05-APR-1999; 99US-0127728.
 XX PR 30-MAR-2000; 2000US-0540763.
 XX PA (CURA-) CURAGEN CORP.
 XX P1 Shimkets RA, Leach M;
 XX DR WPI; 2000-602362/57.
 XX DR N-PSDB; AAC77446.
 XX PT Novel nucleic acids and peptides derived from open reading frame X,
 XX PT useful for treating e.g. cancers, proliferative disorders,
 XX PS neurodegenerative disorders and cardiovascular disease -
 XX Claim 11; Page 5181-5182; 5507tp; English.
 XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORF open reading frames 1 to 3161. The ORF
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antiproliferative; antiparinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;
 CC antithyroid; and antiaiemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORF-associated disorder. The
 CC nucleic acids can be used to express ORF proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hyperetation, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX XX
 SQ Sequence 263 AA;
 Query Match 99.4%; Score 1379; DB 21; Length 263;
 Best Local Similarity 99.2%; Pred. No. 1.4e-118;
 Matches 260; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 121 GGADEVQYFAPANALPARSEAAVQPIGHSORVRNNSKEKDLGTLGYLITMWII 180
 DB 122 GGADEVQYFAPANALPARSEAAVQPIGHSORVRNNSKEKDLGTLGYLITMWII 181
 QY 181 IAIAGITIGSYKRGKDKKEQHDQKVEREMORITPLSAFNTPTCEIIVDEKTVVHTS 240
 DB 182 IAIAGITIGSYKRGKDKKEQHDQKVEREMORITPLSAFNTPTCEIIVDEKTVVHTS 241
 QY 241 QTPVDPQEGSTPLMGQAGTGA 262
 DB 242 QTPVDPQEGSTPLMGQAGTGA 263
 RESULT 6
 AAM93748
 ID AAM93748 standard; Protein; 263 AA.
 XX AC AAM93748;
 XX DT 06-NOV-2001 (first entry)
 XX DE Human polypeptide, SEQ ID NO: 3727.
 XX KW Human; full length cDNA; cDNA synthesis; oligo-capping.
 XX OS Homo sapiens.
 XX PN EP130094-A2.
 XX PD 05-SEP-2001.
 XX PF 07-JUL-2000; 2000EP-0114089.
 XX PR 08-JUL-1999; 99JP-0194486.
 XX PR 11-JAN-2000; 2000JP-0118774.
 XX PR 02-MAY-2000; 2000JP-0183765.
 XX PA (HELI-) HELIX RES INST.
 XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX DR WPI: 2001-524255/58.
 XX DR N-PSDB; AAK94700.
 XX PT 830 Primers useful for synthesizing full length cDNA clones and their
 XX PT use in genetic manipulation -
 XX PS Claim 8; SEQ ID NO 3727; 1380bp + sequence listing; English.
 XX CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5' - and 3' -ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesized by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 XX XX
 SQ Sequence 263 AA;
 Query Match 99.0%; Score 1374; DB 22; Length 263;
 Best Local Similarity 99.2%; Pred. No. 4.1e-118;
 Matches 260; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 61 SGAGNHSYCRNPDEDPGRGWCYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASE 120
 DB 62 SGAGNHSYCRNPDEDPGRGWCYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASE 121
 QY 121 GPGADEVQVFAPANALPARSEAAAVQPIVIGISQVRNMSKEKDLGTLGYVLGITMMVI 180
 DB 122 GPGADEVQVFAPANALPARSEAAAVQPIVIGISQVRNMSKEKDLGTLGYVLGITMMVI 181
 QY 181 IAIAGGIIILGYSYKRGKDLKEQHDQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTS 240
 DB 182 IAIAGGIIILGYSYKRGKDLKEQHDQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTS 241
 QY 241 QTPVPDQSGSTPLMGQAGTPGA 262
 DB 242 QTPVPDQSGSTPLMGQAGTPGA 263

RESULT 7
 AAY05220
 ID AAY05220 standard; Protein; 286 AA.
 XX AAY05220;
 AC AAY05220;
 DT 17-JUN-1999 (first entry)
 DE Kringel protein sequence.
 XX Kringel1
 KW Kringel1; cancer; inflammation; autoimmunity; rheumatoid arthritis; AIDS;
 KW CNS inflammation; cerebellar degeneration; Alzheimer's disease; asthma;
 KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
 KW neurological abnormality; ischaemia reperfusion injury; ischaemic injury;
 KW cardiovascular disease; kidney disease; liver disease; aplastic anaemia;
 KW myocardial infarction; hypertension; allergy; infection;
 KW myelodysplastic syndrome; haematologic abnormality; diagnosis; therapy;
 KW male pattern baldness.
 XX Homo sapiens.
 OS Homo sapiens.
 XX MO9911788-A1.
 PN 11-MAR-1999.
 PD 11-MAR-1999.
 XX 02-SEP-1998; 98WO-US18270.
 PF 02-SEP-1998; 98US-0144889.
 PR 01-SEP-1998; 98US-0144889.
 PR 02-SEP-1997; 97US-0056032.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA Albone EF, Kikly KK;
 PI WPI; 1999-214707/18.
 DR N-PSDB; AAX28355.
 XX New kringel polypeptides and polynucleotides
 PS Claim 14; Page 33; 42pp; English.

This sequence is a Kringel1 polypeptide of the invention.
 The Kringel1 polypeptides (I) are used to screen for agonists and antagonists. Agonists are used to treat subjects in need of enhanced activity or expression of (I). Antagonists are used to treat subjects having need to inhibit the activity or expression of (I). The methods can be used to treat conditions such as cancer, inflammation, autoimmunity, allergy, asthma, rheumatoid arthritis, CNS inflammation, cerebellar degeneration, Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis, head injury damage and other neurological abnormalities, ischaemia reperfusion injury, cardiovascular disease, kidney disease, liver disease, ischaemic injury, myocardial infarction, hypertension, AIDS, myelodysplastic syndromes and other haematologic abnormalities, aplastic anaemia, male pattern baldness, and bacterial, fungal, protozoan and viral infections. The

CC kringel1 polypeptides may also be used to generate antibodies.
 CC Determining the presence or absence of mutations in, and analysing for
 CC the presence or absence of expression of, kringel polynucleotides can be
 CC used to diagnose a disease or susceptibility to a disease related to
 CC expression or activity of kringel1 proteins. The polynucleotides may also
 CC be used for chromosome identification, and mapping.

SQ Sequence 286 AA;
 Query Match 83.8%; Score 1162.5; DB 20; Length 286;
 Best local similarity 87.3%; Pred. No. 1,3e-98;
 Matches 226; Conservative 5; Mismatches 27; Indels 1; Gaps 1;

QY 1 LLAWQAFVLSNMLLAEAYGSGCFWMDNGHLYREDQTSAPGLRCILNWLDAQSGLASAPV 60
 DB 2 LLAWQAFVLSNMLLAEAYGSGCFWMDNGHLYREDQTSAPGLRCILNWLDAQSGLASAPV 61
 QY 61 SGAGNHSYCRNPDEDPGRGWCYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEAS 119
 DB 62 SGAGNHSYCRNPDEDPGRGWCYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEAS 121
 QY 120 EGPDADEVQVFAPANALPARSEAAAVQPIVIGISQVRNMSKEKDLGTLGYVLGITMMVI 179
 DB 122 EGPDADEVQVFAPANALPARSEAAAVQPIVIGISQVRNMSKEKDLGTLGYVLGITMMVI 181
 QY 180 IAIAGGIIILGYSYKRGKDLKEQHDQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHT 239
 DB 182 IAIAGGIIILGYSYKRGKDLKEQHDQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHT 241
 QY 240 SQTVPDQSGSTPLMGQAG 258
 DB 242 SQTVPDQSGSTPLMGQAG 260

RESULT 8
 ABB37905
 ID ABB37905 standard; Peptide; 66 AA.
 XX ABB37905;
 AC ABB37905;
 DT 04-FEB-2002 (first entry)
 DE Peptide #5411 encoded by human foetal liver single exon probe.
 KW Human; foetal liver; gene expression; single exon nucleic acid probe.
 OS Homo sapiens.
 XX MO200157277-A2.
 PN 09-AUG-2001.
 PD 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US00669.
 PF 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-483447/52.
 DR Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -
 XX Claim 27; SEQ ID NO 30540; 639pp + sequence listing; English.

CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 66 AA;

Query Match 23.2%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 3.5e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 TTSQLPFTTEIOEASGPGADEVQVFPANALPARSEAAAVQVIGISQVRMNSKEK 162
Db 1 TTSQLPFTTEIOEASGPGADEVQVFPANALPARSEAAAVQVIGISQVRMNSKEK 60

QY 163 KDLGTL 168
Db 61 KDLGTL 66

RESULT 9
AAB23159
ID AAB23159 standard; Protein; 66 AA.
XX
AC AAB23159;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #5158 encoded by probe for measuring heart cell gene expression.
XX
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX
PS Claim 15; SEQ ID No 24929; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA41353-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the

CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 66 AA;

Query Match 23.2%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 3.5e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 TTSQLPFTTEIOEASGPGADEVQVFPANALPARSEAAAVQVIGISQVRMNSKEK 162
Db 1 TTSQLPFTTEIOEASGPGADEVQVFPANALPARSEAAAVQVIGISQVRMNSKEK 60

QY 163 KDLGTL 168
Db 61 KDLGTL 66

RESULT 10
AAMS5837
ID AAMS5837 standard; Protein; 66 AA.
XX
AC AAMS5837;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30642.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
PS Example 4; SEQ ID NO: 30642; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 66 AA;

Query Match 23.2%; Score 322; DB 22; Length 66;

Best Local Similarity 100.0%; Pred. No. 3.5e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 TTSQLPATTTEIOASBGPADDEVQVFAPANALPARSEAAAVQPVIGISQVRNNSKEX 162
DB 1 TTSQLPATTTEIOASBGPADDEVQVFAPANALPARSEAAAVQPVIGISQVRNNSKEX 60
QY 163 KDLGTL 168
DB 61 KDLGTL 66

RESULT 11

ID AAM71037 standard; Protein; 66 AA.

AC AAM71037;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31343.

KM Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukemia; lymphoma; myeloma.

OS Homo sapiens.

PN W0200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00668.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-488900/53.

Human genome-derived single exon nucleic acid probes useful for

analyzing gene expression in human bone marrow -

Example 4; SEQ ID NO: 31343; 658bp + Sequence Listing; English.

The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human bone marrow. They can be used to measure gene expression in bone marrow samples, which may enable the improved diagnosis and treatment of cancers such as lymphoma, leukemia and myeloma. The present sequence is a protein encoded by one of the probes of the invention.

Sequence 66 AA;

Query Match 23.2%; Score 322; DB 22; Length 66;

Best Local Similarity 100.0%; Pred. No. 3.5e-22;

Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 TTSQLPATTTEIOASBGPADDEVQVFAPANALPARSEAAAVQPVIGISQVRNNSKEX 162

DB 1 TTSQLPATTTEIOASBGPADDEVQVFAPANALPARSEAAAVQPVIGISQVRNNSKEX 60

QY 163 KDLGTL 168

DB 61 KDLGTL 66

RESULT 12

ID AAM18800 standard; Protein; 66 AA.

AC AAM18800;

DT 12-OCT-2001 (first entry)

DE Peptide #5234 encoded by probe for measuring cervical gene expression.

KM Probe; human; microarray; gene expression; cervical epithelial cell;

KW cervical cancer.

OS Homo sapiens.

PN W0200157278-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00670.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-488901/53.

Human genome-derived single exon nucleic acid probes useful for

analyzing gene expression in human cervical epithelial cells -

Claim 27; SEQ ID NO 23626; 487bp; English.

The present invention relates to human single exon nucleic acid probes (SENP: see AAI10068-AI28459). The present sequence is a peptide encoded by one such probe. The SENPs are derived from human HeLa cells. The SENPs can be used to produce a single exon microarray, which can be used for measuring human gene expression in a sample derived from human cervical epithelial cells. By measuring gene expression, the probes are therefore useful in grading and/or staging of diseases of the cervix, notably cervical cancer.

Note: The sequence data for this patent did not form part of the printed CC specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 66 AA;

Query Match 23.2%; Score 322; DB 22; Length 66;

Best Local Similarity 100.0%; Pred. No. 3.5e-22;

Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 TTSQLPATTTEIOASBGPADDEVQVFAPANALPARSEAAAVQPVIGISQVRNNSKEX 162

DB 1 TTSQLPATTTEIOASBGPADDEVQVFAPANALPARSEAAAVQPVIGISQVRNNSKEX 60

QY 163 KDLGTL 168

DB 61 KDLGTL 66

RESULT 13

ID AAM31314 standard; Protein; 66 AA.

AC AAM31314;

XX 17-OCT-2001 (first entry)
XX Peptide #5351 encoded by probe for measuring placental gene expression.
DE Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX Homo sapiens.
OS
XX WO200157272-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US00663.
PF
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488897/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
XX Claim 27; SEQ ID No 31583; 654bp; English.
PS
XX The present invention relates to single exon nucleic acid probes (SENPs;
CC see A131315-A157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
XX Sequence 66 AA;
SQ
Query Match 23.2%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 3.5e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 103 TTGQALPATTETIOBAGSEGGADVEQVAPANALPARSEAAAVQPVYIGISORVRMSKEX 162
DB 1 TTSQALPATTETIOBAGSEGGADVEQVAPANALPARSEAAAVQPVYIGISORVRMSKEX 60
QY 163 KDLGTL 168
DB 61 KDLGTL 66
RESULT 14
ABG40828
ID ABG40828 standard; Peptide; 66 AA.
XX
AC ABG40828;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human peptide encoded by genome-derived single exon probe SEQ ID 30493.
XX
XX Human: single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hereditary-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
XX Homo sapiens.
OS
XX WO200186003-A2.
PN
XX 15-NOV-2001.
PD
XX 30-JAN-2001; 2001WO-US00665.
PF
XX 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2002-114183/15.
DR
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -
XX
XX Claim 27; SEQ ID No 30493; 634bp; English.
PS
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes. The novel set of probes which hybridize at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridization of detectably
CC labeled nucleic acids from eukaryotic lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridization to a single exon
CC microarray having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung disease
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hereditary-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a peptide/protein
CC encoded by a single exon probe of the invention.
XX Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from Wipo at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 66 AA;
SQ

Query Match 23.2%; Score 322; DB 23; Length 66;
 Best Local Similarity 100.0%; Pred. No. 3.5e-22;
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 TTSQLPAPFTTEIOBESGPGADEVQVFAPANMLPARSEMAAVQVPIGIGSRVNNSKER 162
 DB 1 TTSQLPAPFTTEIOBESGPGADEVQVFAPANMLPARSEMAAVQVPIGIGSRVNNSKER 60

QY 163 KDLGTL 168
 DB 61 KDLGTL 66

RESULT 15
 ID AAY12615
 XX AAY12615 standard; Protein: 56 AA.
 AC AAY12615;
 XX
 DT 22-JUN-1999 (first entry)
 XX
 DE Human 5' EST secreted protein SEQ ID NO: 280 from WO 9906553.
 XX
 KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide;
 KW upstream regulatory sequence; cytokine activity; cell proliferation;
 KW differentiation; haematopoiesis regulation; tissue growth regulation;
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; antiinflammatory; tumour inhibition; antitumour.
 XX
 OS Homo sapiens.
 XX
 PN WO9906553-A2.
 XX
 PD 11-FEB-1999.
 XX
 PF 31-JUL-1998; 98WO-IB01237.
 XX
 PR 01-AUG-1997; 97US-0905051.
 XX
 PA (GEST) GENSET.
 XX
 PI Duclert A, Dumas M,ne Edwards J, Lacroix B;
 XX
 DR MPI: 1999-153783/13.
 DR N-PSDB; AAX41473.
 XX
 PT New nucleic acids encoding human secreted proteins - obtained from
 PT cDNA libraries derived from umbilical cord, lymph ganglia,
 XX lymphocytes and placental tissue
 XX
 PS Claim 34; Page 376; 411p; English.
 XX
 CC AAX41379 to AAX41526 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAY12521 to
 CC AAY12668, respectively. The proteins given represent the signal peptide
 CC and an N-terminal fragment of a secreted protein. The nucleic acid
 CC sequences can be used for producing secreted human gene products. They
 CC can also be used to develop products for diagnosis and therapy. The
 CC proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, haematopoiesis regulating
 CC activity, tissue growth regulating activity, reproductive hormone
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
 CC thrombolytic activity, receptor/ligand activity, antiinflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.
 XX
 SQ Sequence + 56 AA;

Query Match 21.7%; Score 301; DB 20; Length 56;
 Best Local Similarity 98.2%; Pred. No. 2.4e-20;
 Matches 54; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILAVYQAEFLVSNMLLAEBAYSGGCGFMDNGHLYREDQTSAPGIRCLINWLDQSG 55
 DB 2 ILAVYQAEFLVSNMLLAEBAYSGGCGFMDNGHLYREDQTSAPGIRCLINWLDQNG 56

Search completed: April 7, 2003, 08:42:30
 Job time : 38 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 08:43:33 ; Search time 16 Seconds

(without alignments)
1001.103 Million cell updates/sec

Title: US-10-057-951-2_COPY_2_263

Sequence: 1 LIAWVQAFIVSNMLLAAYG.....PVDQSGSRPLMGQACTGCA 262

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lasting first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCR_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 1388 | 100.0 | 263 | 9 | US-10-102-704-2 |
| 2 | 1388 | 100.0 | 263 | 9 | US-10-057-951-2 |
| 3 | 1388 | 100.0 | 263 | 10 | US-09-084-491A-2 |
| 4 | 322 | 23.2 | 66 | 10 | US-09-864-761-38457 |
| 5 | 154.5 | 11.1 | 527 | 9 | US-09-987-457-18 |
| 6 | 154.5 | 11.1 | 527 | 9 | US-09-987-455-19 |
| 7 | 154.5 | 11.1 | 562 | 9 | US-09-974-288-145 |
| 8 | 154.5 | 11.1 | 562 | 10 | US-09-969-271-7 |
| 9 | 142 | 10.2 | 160 | 9 | US-09-335-325-35 |
| 10 | 142 | 10.2 | 160 | 9 | US-10-131-241-35 |
| 11 | 142 | 10.2 | 160 | 10 | US-09-761-120-35 |
| 12 | 142 | 10.2 | 250 | 9 | US-09-335-325-30 |
| 13 | 142 | 10.2 | 250 | 10 | US-10-131-241-30 |
| 14 | 142 | 10.2 | 250 | 10 | US-09-761-120-30 |
| 15 | 142 | 10.2 | 260 | 9 | US-10-131-241-61 |
| 16 | 142 | 10.2 | 339 | 9 | US-09-335-325-3 |
| 17 | 142 | 10.2 | 339 | 9 | US-10-131-241-3 |
| 18 | 142 | 10.2 | 339 | 10 | US-09-761-120-3 |
| 19 | 142 | 10.2 | 339 | 10 | US-09-761-120-3 |

| | | | | | | |
|----|-----|------|-----|----|------------------|-------------------|
| 20 | 142 | 10.2 | 352 | 9 | US-09-335-325-40 | Sequence 40, Appl |
| 21 | 142 | 10.2 | 352 | 9 | US-10-131-241-40 | Sequence 40, Appl |
| 22 | 142 | 10.2 | 352 | 10 | US-09-761-120-40 | Sequence 40, Appl |
| 23 | 142 | 10.2 | 368 | 10 | US-09-761-120-42 | Sequence 42, Appl |
| 24 | 142 | 10.2 | 378 | 9 | US-09-335-325-42 | Sequence 42, Appl |
| 25 | 142 | 10.2 | 378 | 9 | US-10-131-241-42 | Sequence 42, Appl |
| 26 | 142 | 10.2 | 378 | 10 | US-09-873-676-1 | Sequence 1, Appl |
| 27 | 142 | 10.2 | 458 | 10 | US-09-946-893-4 | Sequence 4, Appl |
| 28 | 142 | 10.2 | 569 | 10 | US-09-946-893-5 | Sequence 5, Appl |
| 29 | 142 | 10.2 | 571 | 10 | US-09-946-893-8 | Sequence 8, Appl |
| 30 | 142 | 10.2 | 576 | 10 | US-09-946-893-6 | Sequence 6, Appl |
| 31 | 142 | 10.2 | 791 | 9 | US-09-967-386-1 | Sequence 1, Appl |
| 32 | 142 | 10.2 | 810 | 10 | US-09-946-893-2 | Sequence 2, Appl |
| 33 | 138 | 9.9 | 79 | 9 | US-09-335-325-8 | Sequence 8, Appl |
| 34 | 138 | 9.9 | 79 | 9 | US-10-131-241-8 | Sequence 8, Appl |
| 35 | 138 | 9.9 | 79 | 10 | US-09-753-064-2 | Sequence 2, Appl |
| 36 | 138 | 9.9 | 79 | 10 | US-09-761-120-8 | Sequence 8, Appl |
| 37 | 135 | 9.7 | 377 | 9 | US-09-87-455-8 | Sequence 8, Appl |
| 38 | 133 | 9.6 | 326 | 9 | US-10-057-951-3 | Sequence 3, Appl |
| 39 | 133 | 9.6 | 354 | 9 | US-09-987-457-10 | Sequence 10, Appl |
| 40 | 133 | 9.6 | 354 | 9 | US-09-987-455-11 | Sequence 11, Appl |
| 41 | 133 | 9.6 | 372 | 9 | US-10-102-704-3 | Sequence 3, Appl |
| 42 | 133 | 9.6 | 372 | 10 | US-09-084-491A-3 | Sequence 3, Appl |
| 43 | 132 | 9.5 | 411 | 10 | US-09-880-503-3 | Sequence 3, Appl |
| 44 | 132 | 9.5 | 431 | 9 | US-10-076-421-2 | Sequence 2, Appl |
| 45 | 132 | 9.5 | 431 | 10 | US-09-264-468B-1 | Sequence 1, Appl |

ALIGNMENTS

RESULT 1
US-10-102-704-2
Sequence 2, Application US/10102704
Patent No. US20020164768A1
GENERAL INFORMATION:
APPLICANT: Moore et al.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protein
FILE REFERENCE: PE378C1
CURRENT APPLICATION NUMBER: US/10/102,704
CURRENT FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/084,491
PRIOR FILING DATE: 1998-05-27
PRIOR APPLICATION NUMBER: 60/048,000
PRIOR FILING DATE: 1997-05-28
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-10-102-704-2

| Query Match | Score | 1388 | DB 9 | Length | 263 |
|-----------------------|--------|---|----------|------------|-----|
| Best Local Similarity | 100.0% | Pred. No. | 5,9e-119 | Indels | 0 |
| Matches | 262 | Conservative | 0 | Mismatches | 0 |
| Qy | 1 | LIAWVQAFIVSNMLLAAYGSGCFWNGHLYREDQTSAPAGRLCLNWDAGSLASAPV | 60 | | |
| Db | 2 | LIAWVQAFIVSNMLLAAYGSGCFWNGHLYREDQTSAPAGRLCLNWDAGSLASAPV | 61 | | |
| Qy | 61 | SGAGNYSYRNDEDEPRGPMVCVSGEAVPEKRCPELRCPETTSQALPAFTTEIOEASE | 120 | | |
| Db | 62 | SGAGNYSYRNDEDEPRGPMVCVSGEAVPEKRCPELRCPETTSQALPAFTTEIOEASE | 121 | | |
| Qy | 121 | GPADAEVQFAPANLAPARSEAAAVPYIGISQVRNNSREKDLGTLGVVGTITMWTI | 180 | | |
| Db | 122 | GPADAEVQFAPANLAPARSEAAAVPYIGISQVRNNSREKDLGTLGVVGTITMWTI | 181 | | |
| Qy | 181 | IATGAGIIGYSYKSGKDLKEQHDQVCEREMQITPLSAFTNPCEIIVDEKTVVHTS | 240 | | |
| Db | 182 | IATGAGIIGYSYKSGKDLKEQHDQVCEREMQITPLSAFTNPCEIIVDEKTVVHTS | 241 | | |

QY 241 QTPVDPQEGSTPLMGQAGTPGA 262
DB 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 2

US-10-057-951-2
Sequence 2, Application US/10057951
Patent No. US20020177213A1
GENERAL INFORMATION:
APPLICANT: Moore et al.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
FILE REFERENCE: PF378P1
CURRENT APPLICATION NUMBER: US/10/057,951
PRIOR FILING DATE: 2002-01-29
PRIOR APPLICATION NUMBER: US 09/411,977
PRIOR FILING DATE: 1999-10-04
PRIOR APPLICATION NUMBER: US 09/084,491
PRIOR FILING DATE: 1998-05-27
PRIOR APPLICATION NUMBER: US 60/048,000
PRIOR FILING DATE: 1997-05-28
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO: 2
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-10-057-951-2

Query Match

Best Local Similarity 100.0%; Score 1388; DB 9; Length 263;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAWQAFVLSNMLAEAYSGGCFMNDGHLVREDQTSPPAGLRCLNMLDAQSGIASAPV 60
DB 2 LLAWQAFVLSNMLAEAYSGGCFMNDGHLVREDQTSPPAGLRCLNMLDAQSGIASAPV 61
QY 61 SGAGNSHYCRNPDEDPKRCVSGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIOEASE 120
DB 62 SGAGNSHYCRNPDEDPKRCVSGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIOEASE 121
QY 121 GPGADEVQVAPANALPARSEAAVQPIVIGISQVRNNSKEKDLGTLGVIGITMVI 180
DB 122 GPGADEVQVAPANALPARSEAAVQPIVIGISQVRNNSKEKDLGTLGVIGITMVI 181
QY 181 IAIAGIILIGSYKRGKDLKEQHDQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTS 240
DB 182 IAIAGIILIGSYKRGKDLKEQHDQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTS 241
QY 241 QTPVDPQEGSTPLMGQAGTPGA 262
DB 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 3

US-09-084-491A-2
Sequence 2, Application US/09084491A
Patent No. US20020061576A1
GENERAL INFORMATION:
APPLICANT: MOORE, PAUL A.
APPLICANT: RUBEN, STEVEN M.
APPLICANT: BENNER, REINHARD
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/084,491A
FILING DATE: 27-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PP378
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-084-491A-2

Query Match 100.0%; Score 1388; DB 10; Length 263;
Best Local Similarity 100.0%; Pred. No. 5,9e-119; Indels 0; Gaps 0;
Matches 262; Conservative 0; Mismatches 0;

QY 1 LLAWQAFVLSNMLAEAYSGGCFMNDGHLVREDQTSPPAGLRCLNMLDAQSGIASAPV 60
DB 2 LLAWQAFVLSNMLAEAYSGGCFMNDGHLVREDQTSPPAGLRCLNMLDAQSGIASAPV 61
QY 61 SGAGNSHYCRNPDEDPKRCVSGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIOEASE 120
DB 62 SGAGNSHYCRNPDEDPKRCVSGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIOEASE 121
QY 121 GPGADEVQVAPANALPARSEAAVQPIVIGISQVRNNSKEKDLGTLGVIGITMVI 180
DB 122 GPGADEVQVAPANALPARSEAAVQPIVIGISQVRNNSKEKDLGTLGVIGITMVI 181
QY 181 IAIAGIILIGSYKRGKDLKEQHDQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTS 240
DB 182 IAIAGIILIGSYKRGKDLKEQHDQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTS 241
QY 241 QTPVDPQEGSTPLMGQAGTPGA 262
DB 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 4

US-09-864-761-38457
Sequence 38457, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30

```
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38457
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC002073.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
; OTHER INFORMATION: SWISSPROT HIT: P39881, EVALU 3.40e+00
; OTHER INFORMATION: EST_HUMAN HIT: BE299406.1, EVALU 5.00e-30
US-09-864-761-38457

Query Match
Best Local Similarity 100.0%; Pred.No. 1.7e-22; Length 66;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 TTGQALPFTTETQASGEGADEVOVPAPANALPARSEAAAVQPTGISORVMSKEX 162
Db 1 TTGQALPFTTETQASGEGADEVOVPAPANALPARSEAAAVQPTGISORVMSKEX 60
QY 163 KDLGTL 168
Db 61 KDLGTL 66

RESULT 5
US-09-987-457-18
; Sequence 18, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Rolf-Guenther
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; PRIOR FILING DATE: 2001-11-14
; CURRENT APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
```

```
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens (tpa)
US-09-987-457-18

Query Match
Best Local Similarity 11.1%; Score 154.5; DB 9; Length 527;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 24 CFWDNCHLYREDOQSPARGLCLWMDQSLASAPVS-----GAGNHSYCRNPDE 74
Db 92 CYEDQGISYRGWTASGAECTNW--NSSALQKPYSGRRPDALRLGNNHYCRNPDR 149
QY 75 DPERGWCYVSGEAGVPEKRPCEDLRCPETTS 105
Db 150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 178

RESULT 6
US-09-987-455-19
; Sequence 19, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Goetz, Rolf-Guenther
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for large Scale Production of Recombinant
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-987-455-19

Query Match
Best Local Similarity 11.1%; Score 154.5; DB 9; Length 527;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 24 CFWDNCHLYREDOQSPARGLCLWMDQSLASAPVS-----GAGNHSYCRNPDE 74
Db 92 CYEDQGISYRGWTASGAECTNW--NSSALQKPYSGRRPDALRLGNNHYCRNPDR 149
QY 75 DPERGWCYVSGEAGVPEKRPCEDLRCPETTS 105
Db 150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 178

RESULT 7
US-09-974-298-145
; Sequence 145, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hui-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
```

```
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 145
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CD1
US-09-974-298-145
```

```
Query Match 11.1%; Score 154.5; DB 9; Length 562;
Best Local Similarity 39.6%; Pred. No. 5.1e-06;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
```

```
Qy 24 CFWDNGHLYREDQSPAPGLRCINWMDAOSGLASAPVS-----GAGNHSYCRNPDE 74
Db 127 CLEDGISTYRGWTASTAGACTNW--NSALAKQKYSGRDPDAIRLIGLNHNVCNRPDR 184
```

```
Qy 75 DPRGWCYSGEAGVPEKRPCEDLRCPEETS 105
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213
```

```
RESULT 8
US-09-969-271-7
; Sequence 7, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT FILING DATE: 2001-10-01
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: GB 0025473.0
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-271-7
```

```
Query Match 11.1%; Score 154.5; DB 10; Length 562;
Best Local Similarity 39.6%; Pred. No. 5.1e-06;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
```

```
Qy 24 CFWDNGHLYREDQSPAPGLRCINWMDAOSGLASAPVS-----GAGNHSYCRNPDE 74
Db 127 CLEDGISTYRGWTASTAGACTNW--NSALAKQKYSGRDPDAIRLIGLNHNVCNRPDR 184
```

```
Qy 75 DPRGWCYSGEAGVPEKRPCEDLRCPEETS 105
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213
```

```
RESULT 9
US-09-335-325-35
; Sequence 35, Application US/09335325
; Patent No. US20020164717A1
; GENERAL INFORMATION:
; APPLICANT: Polkman, M. Judah
; O'Reilly, Michael
; Cao, Yihai
; Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
```

```
STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/335,325
FILING DATE: 17-Jun-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/612,788
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Warren, William L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 05213-0126
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 160 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: KI-2
SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-335-325-35
```

```
Query Match 10.2%; Score 142; DB 9; Length 160;
Best Local Similarity 37.2%; Pred. No. 1.4e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;
```

```
Qy 24 CFWDNGHLYREDQSPAPGLRCINWMDA-----OSGLASAPVSGAGNHSYCRNPDEPRG 78
Db 1 CKTGNKNGVGTWSTKRNKNGITCQKWSSTSPHRRPSPATHPSEGL-BENYCRNPDPNDPOG 59
```

```
Qy 79 PWCYSGEAGVPEKRPCEDLRCPE 102
Db 60 PWCYTD---PEKRYDYCDILECEE 81
```

```
RESULT 10
US-10-131-241-35
; Sequence 35, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
```

Page 5

STREET: 191 Peachtree Street, 37th Floor

; SOFTWARE: PatentIn version 3.1

Mon Apr 7 10:22:35 2003

us-10-057-951-2_copy_2_263.rapb

Page 6

SEQ ID NO 30
LENGTH: 250
TYPE: PRT
ORGANISM: Homo sapiens
US-10-131-241-30

Query Match 10.2%; Score 142; DB 9; Length 250;
Best Local Similarity 37.2%; Pred. No. 2.5e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 24 CFWDNGHLYREDQTSAPGLRCLNWLDA-----QSGLASAPVSGAGNHSYCRNPDEDPFG 78
DB 1 CKTGNGKXVYRGWMSKTKNGGITCQKMSSTSPHRPRFSPATHPSBGL-ENNYCRNPNDPFG 59
QY 79 PWCYVSGEAGVPEKR--PCEDLRCP 102
DB 60 PWCYTTD---PEKRYDYCDILECEE 81

RESULT 14
US-09-761-120-30
Sequence 30 Application US/09761120
Patent No US20020037847A1
GENERAL INFORMATION:
APPLICANT: Folman, M. Judah
APPLICANT: O'Reilly, Michael
TITLE OF INVENTION: Nucleic Acid Encoding Kringle 1-5 Region Fragments of Plasminogen
FILE REFERENCE: 05940-0151 (43171-252068)
CURRENT APPLICATION NUMBER: US/09/761,120
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 09/309,821
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 08/866,735
PRIOR FILING DATE: 1997-05-30
NUMBER OF SEQ ID NOS: 47
SOFTWARE: PatentIn version 3.0
SEQ ID NO 30
LENGTH: 250
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Kringle 1-3
US-09-761-120-30

Query Match 10.2%; Score 142; DB 10; Length 250;
Best Local Similarity 37.2%; Pred. No. 2.5e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 24 CFWDNGHLYREDQTSAPGLRCLNWLDA-----QSGLASAPVSGAGNHSYCRNPDEDPFG 78
DB 1 CKTGNGKXVYRGWMSKTKNGGITCQKMSSTSPHRPRFSPATHPSBGL-ENNYCRNPNDPFG 59
QY 79 PWCYVSGEAGVPEKR--PCEDLRCP 102
DB 60 PWCYTTD---PEKRYDYCDILECEE 81

RESULT 15
US-10-131-241-61
Sequence 61 Application US/10131241
Publication No. US20030012792A1
GENERAL INFORMATION:
APPLICANT: Holaday, John W.
APPLICANT: Fortier, Anne H.
TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
FILE REFERENCE: 05213-0344 43170-271565
CURRENT APPLICATION NUMBER: US/10/131,241
PRIOR FILING DATE: 2002-07-22
PRIOR APPLICATION NUMBER: US 09/413,049
PRIOR FILING DATE: 1999-10-06
PRIOR APPLICATION NUMBER: US 09/316,802

PRIOR FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: US 60/086,586
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.1
SEQ ID NO 61
LENGTH: 260
TYPE: PRT
ORGANISM: Homo sapiens
US-10-131-241-61

Query Match 10.2%; Score 142; DB 9; Length 260;
Best Local Similarity 37.2%; Pred. No. 2.6e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 24 CFWDNGHLYREDQTSAPGLRCLNWLDA-----QSGLASAPVSGAGNHSYCRNPDEDPFG 78
DB 6 CKTGNGKXVYRGWMSKTKNGGITCQKMSSTSPHRPRFSPATHPSBGL-ENNYCRNPNDPFG 64
QY 79 PWCYVSGEAGVPEKR--PCEDLRCP 102
DB 65 PWCYTTD---PEKRYDYCDILECEE 86

Search completed: April 7, 2003, 08:50:49
Job time: 17 secs

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 08:42:37 ; Search time 328 Seconds

(without alignments)
515,000 Million cell updates/sec

Title: US-10-057-951-2_COPY_2_263

Perfect score: 1388
Sequence: 1 LLAWQAFIVSNMLLAAYG.....PVDQEGSTPLMGQAPGA 262

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:*

1: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US09_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US10_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US10_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US10_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 1388 | 100.0 | 263 | 1 | PCT-US00-27239-2 |
| 2 | 1388 | 100.0 | 263 | 1 | PCT-US02-05301-164 |
| 3 | 1388 | 100.0 | 263 | 1 | PCT-US98-10728-2 |
| 4 | 1388 | 100.0 | 263 | 14 | US-09-084-8914-2 |
| 5 | 1388 | 100.0 | 263 | 15 | US-09-144-889-2 |
| 6 | 1388 | 100.0 | 263 | 24 | US-10-057-951-2 |

| | | | | | | |
|----|------|-------|-----|----|---------------------|--------------------|
| 7 | 1388 | 100.0 | 263 | 25 | US-10-102-704-2 | Sequence 2, Appl1 |
| 8 | 1388 | 100.0 | 263 | 27 | US-60-048-000-2 | Sequence 2, Appl1 |
| 9 | 1385 | 99.8 | 263 | 21 | US-09-791-537-39562 | Sequence 39562, A |
| 10 | 1385 | 99.8 | 263 | 23 | US-09-927-796-44 | Sequence 44, Appl1 |
| 11 | 1385 | 99.8 | 263 | 26 | US-10-210-951-44 | Sequence 44, Appl1 |
| 12 | 1385 | 99.8 | 263 | 26 | US-10-211-858-44 | Sequence 44, Appl1 |
| 13 | 1385 | 99.8 | 263 | 26 | US-10-211-884-44 | Sequence 44, Appl1 |
| 14 | 1385 | 99.8 | 263 | 26 | US-60-230-435-1590 | Sequence 1590, Ap |
| 15 | 1385 | 99.8 | 263 | 27 | US-60-389-987-244 | Sequence 244, App |
| 16 | 1385 | 99.8 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 17 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 18 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 19 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 20 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 21 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 22 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 23 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 24 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 25 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 26 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 27 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 28 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 29 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 30 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 31 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 32 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 33 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 34 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 35 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 36 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 37 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 38 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 39 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 40 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 41 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 42 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 43 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 44 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 45 | 1374 | 99.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |

ALIGNMENTS

RESULT 1
PCT-US00-27239-2
Sequence 2, Application PC/TUS0027239
GENERAL INFORMATION:
APPLICANT: HUMAN GENOME SCIENCES, INC.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
FILE REFERENCE: P378PCT2
CURRENT APPLICATION NUMBER: PCT/US00/27239
CURRENT FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 09/411,977
PRIOR FILING DATE: 1999-10-04
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US00-27239-2

Query Match 100.0%; Score 1388; DB 1; Length 263;
Best local similarity 100.0%; Pred. No. 1.8e-128;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAWQAFIVSNMLLAAYGSGGCFDNGHLYREQOTSPAPGLCLNMLDAOSGLSARV 60
DB 2 LLAWQAFIVSNMLLAAYGSGGCFDNGHLYREQOTSPAPGLCLNMLDAOSGLSARV 61
QY 61 SGAQHNSTYRNPDDEPRGPFCTVSGAGVPEKRPEDRCRETTSSQALPFTTIOEASE 120

Db 62 SGAGNHSYCRNPDEDPGPMCVYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASE 121
Qy 121 GGADDEVQVFPANALPARSEAAAVQVIGISQVRNMSKEKDLGLTGLVIGITMVYI 180
Db 122 GGADDEVQVFPANALPARSEAAAVQVIGISQVRNMSKEKDLGLTGLVIGITMVYI 181
Qy 181 IAIAGGIIIGSYKRGKDLKEQHDQVCEREMQRTITPLSAFTNPCEIYDEKTVVHTS 240
Db 182 IAIAGGIIIGSYKRGKDLKEQHDQVCEREMQRTITPLSAFTNPCEIYDEKTVVHTS 241
Qy 241 QTPVDPQEGSTPLMGOAGTRGA 262
Db 242 QTPVDPQEGSTPLMGOAGTRGA 263

RESULT 2
PCT-US02-05301-164
Sequence 164, Application PC/TUS0205301
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: 70 Human Secreted Proteins
FILE REFERENCE: PS736PCT
CURRENT APPLICATION NUMBER: PCT/US02/05301
PRIOR FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: US 60/304,417
PRIOR FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: US 60/270,625
NUMBER OF SEQ ID NOS: 340
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 164
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-05301-164

Query Match 100.0%; Score 1388; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 1,8e-128;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIAWQAFIVSNMLLAAYSGGCFWMDNGHLTYREDQTSAPAGLRCLNMLDAOSGLASAPV 60
Db 2 LIAWQAFIVSNMLLAAYSGGCFWMDNGHLTYREDQTSAPAGLRCLNMLDAOSGLASAPV 61
Qy 61 SGAGNHSYCRNPDEDPGPMCVYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASE 120
Db 62 SGAGNHSYCRNPDEDPGPMCVYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASE 121
Qy 121 GGADDEVQVFPANALPARSEAAAVQVIGISQVRNMSKEKDLGLTGLVIGITMVYI 180
Db 122 GGADDEVQVFPANALPARSEAAAVQVIGISQVRNMSKEKDLGLTGLVIGITMVYI 181
Qy 181 IAIAGGIIIGSYKRGKDLKEQHDQVCEREMQRTITPLSAFTNPCEIYDEKTVVHTS 240
Db 182 IAIAGGIIIGSYKRGKDLKEQHDQVCEREMQRTITPLSAFTNPCEIYDEKTVVHTS 241
Qy 241 QTPVDPQEGSTPLMGOAGTRGA 262
Db 242 QTPVDPQEGSTPLMGOAGTRGA 263

RESULT 3
PCT-US98-10728-2
Sequence 2, Application PC/TUS9810728
GENERAL INFORMATION:
APPLICANT: BENNER, REINHARD
APPLICANT: MOORE, PAUL
APPLICANT: RUBEN, STEVEN
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE

CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US98/10728
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF379PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US98-10728-2

Query Match 100.0%; Score 1388; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 1,8e-128;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIAWQAFIVSNMLLAAYSGGCFWMDNGHLTYREDQTSAPAGLRCLNMLDAOSGLASAPV 60
Db 2 LIAWQAFIVSNMLLAAYSGGCFWMDNGHLTYREDQTSAPAGLRCLNMLDAOSGLASAPV 61
Qy 61 SGAGNHSYCRNPDEDPGPMCVYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASE 120
Db 62 SGAGNHSYCRNPDEDPGPMCVYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASE 121
Qy 121 GGADDEVQVFPANALPARSEAAAVQVIGISQVRNMSKEKDLGLTGLVIGITMVYI 180
Db 122 GGADDEVQVFPANALPARSEAAAVQVIGISQVRNMSKEKDLGLTGLVIGITMVYI 181
Qy 181 IAIAGGIIIGSYKRGKDLKEQHDQVCEREMQRTITPLSAFTNPCEIYDEKTVVHTS 240
Db 182 IAIAGGIIIGSYKRGKDLKEQHDQVCEREMQRTITPLSAFTNPCEIYDEKTVVHTS 241
Qy 241 QTPVDPQEGSTPLMGOAGTRGA 262
Db 242 QTPVDPQEGSTPLMGOAGTRGA 263

RESULT 4
US-09-084-491A-2
Sequence 2, Application US/09084491A
GENERAL INFORMATION:
APPLICANT: MOORE, PAUL A.
APPLICANT: RUBEN, STEVEN M.
APPLICANT: BENNER, REINHARD
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/084,491A
FILING DATE: 27-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF378
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-084-491A-2

Query Match 100.0%; Score 1388; DB 14; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.8e-128;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAMVOAFIVSNMLAEAYSGGCFMNDGHLTYREDQTSAPAGLRCLNWLDAOSGLASAPV 60
DB 2 LLAMVOAFIVSNMLAEAYSGGCFMNDGHLTYREDQTSAPAGLRCLNWLDAOSGLASAPV 61
QY 61 SGAGNHSYCRNPDEDPGPMCVYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASE 120
DB 62 SGAGNHSYCRNPDEDPGPMCVYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASE 121
QY 121 GPADDEVQVFAFANALPARSEAAAVQVIGISQVRMNSKEKDLGTLGYVIGITMVI 180
DB 122 GPADDEVQVFAFANALPARSEAAAVQVIGISQVRMNSKEKDLGTLGYVIGITMVI 181
QY 181 IAIAGIILIGSYKRGKDLKEOHQKVCEREMORTLPLSAFTNPTCEIYDEKTVVHTS 240
DB 182 IAIAGIILIGSYKRGKDLKEOHQKVCEREMORTLPLSAFTNPTCEIYDEKTVVHTS 241
QY 241 QTPVDPQSGSTPLMGOAGTPGA 262
DB 242 QTPVDPQSGSTPLMGOAGTPGA 263

RESULT 5
US-09-144-889-2

Sequence 2, Application US/09144889B
GENERAL INFORMATION:
APPLICANT: Earl F. Albone
APPLICANT: Kristine K. Kikly
TITLE OF INVENTION: KRINGLE1
FILE REFERENCE: GH-70249
CURRENT APPLICATION NUMBER: US/09/144,889B
CURRENT FILING DATE: 1998-09-01
EARLIER APPLICATION NUMBER: 60/056,032
EARLIER FILING DATE: 1997-09-02
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 2
LENGTH: 263
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-144-889-2

Query Match 100.0%; Score 1388; DB 15; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.8e-128;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAMVOAFIVSNMLAEAYSGGCFMNDGHLTYREDQTSAPAGLRCLNWLDAOSGLASAPV 60
DB 2 LLAMVOAFIVSNMLAEAYSGGCFMNDGHLTYREDQTSAPAGLRCLNWLDAOSGLASAPV 61

QY 61 SGAGNHSYCRNPDEDPGPMCVYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASE 120
DB 62 SGAGNHSYCRNPDEDPGPMCVYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASE 121
QY 121 GPADDEVQVFAFANALPARSEAAAVQVIGISQVRMNSKEKDLGTLGYVIGITMVI 180
DB 122 GPADDEVQVFAFANALPARSEAAAVQVIGISQVRMNSKEKDLGTLGYVIGITMVI 181
QY 181 IAIAGIILIGSYKRGKDLKEOHQKVCEREMORTLPLSAFTNPTCEIYDEKTVVHTS 240
DB 182 IAIAGIILIGSYKRGKDLKEOHQKVCEREMORTLPLSAFTNPTCEIYDEKTVVHTS 241
QY 241 QTPVDPQSGSTPLMGOAGTPGA 262
DB 242 QTPVDPQSGSTPLMGOAGTPGA 263

RESULT 6
US-10-057-951-2

Sequence 2, Application US/10057951
GENERAL INFORMATION:
APPLICANT: Moore et al.
TITLE OF INVENTION: Tissue Plasminogen Activator-like Protease
FILE REFERENCE: PF378P1
CURRENT APPLICATION NUMBER: US/10/057,951
PRIOR FILING DATE: 2002-01-29
PRIOR APPLICATION NUMBER: US 09/411,977
PRIOR FILING DATE: 1999-10-04
PRIOR APPLICATION NUMBER: US 09/084,491
PRIOR FILING DATE: 1998-05-27
PRIOR APPLICATION NUMBER: US 60/048,000
PRIOR FILING DATE: 1997-05-28
NUMBER OF SEQ ID NOS: 31
SOFTWARE: Patent In Ver. 3.1
SEQ ID NO: 2
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-10-057-951-2

Query Match 100.0%; Score 1388; DB 24; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.8e-128;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAMVOAFIVSNMLAEAYSGGCFMNDGHLTYREDQTSAPAGLRCLNWLDAOSGLASAPV 60
DB 2 LLAMVOAFIVSNMLAEAYSGGCFMNDGHLTYREDQTSAPAGLRCLNWLDAOSGLASAPV 61
QY 61 SGAGNHSYCRNPDEDPGPMCVYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASE 120
DB 62 SGAGNHSYCRNPDEDPGPMCVYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASE 121
QY 121 GPADDEVQVFAFANALPARSEAAAVQVIGISQVRMNSKEKDLGTLGYVIGITMVI 180
DB 122 GPADDEVQVFAFANALPARSEAAAVQVIGISQVRMNSKEKDLGTLGYVIGITMVI 181
QY 181 IAIAGIILIGSYKRGKDLKEOHQKVCEREMORTLPLSAFTNPTCEIYDEKTVVHTS 240
DB 182 IAIAGIILIGSYKRGKDLKEOHQKVCEREMORTLPLSAFTNPTCEIYDEKTVVHTS 241
QY 241 QTPVDPQSGSTPLMGOAGTPGA 262
DB 242 QTPVDPQSGSTPLMGOAGTPGA 263

RESULT 7
US-10-102-704-2

Sequence 2, Application US/10102704
GENERAL INFORMATION:
APPLICANT: Moore et al.
TITLE OF INVENTION: Tissue Plasminogen Activator-like Protein
FILE REFERENCE: PF378C1
CURRENT APPLICATION NUMBER: US/10/102,704

```
/ CURRENT FILING DATE: 2002-03-22
/ PRIOR APPLICATION NUMBER: 09/084,491
/ PRIOR FILING DATE: 1998-05-27
/ PRIOR APPLICATION NUMBER: 60/048,000
/ PRIOR FILING DATE: 1997-05-28
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: Patent version 3.1
/ SEQ ID NO: 2
/ LENGTH: 263
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-102-704-2

Query Match
Best Local Similarity 100.0%; Score 1388; DB 25; Length 263;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAWQAFVSNMLLAAYSGGCFWMDNGHLYREDQTSPPAGLRCLNMLDQSGIASAPV 60
DB 2 LLAWQAFVSNMLLAAYSGGCFWMDNGHLYREDQTSPPAGLRCLNMLDQSGIASAPV 61
QY 61 SGAGNHSYCRNPDDPDPGPMCYSGAGVPEKRCEDLRCPETTSQALPAFTTEIOEASE 120
DB 62 SGAGNHSYCRNPDDPDPGPMCYSGAGVPEKRCEDLRCPETTSQALPAFTTEIOEASE 121
QY 121 GGADEVGVFAPANAALPARSEAAAVOPVIGISQVRNNSKEKKDLGTIGYVLGITMVI 180
DB 122 GGADEVGVFAPANAALPARSEAAAVOPVIGISQVRNNSKEKKDLGTIGYVLGITMVI 181
QY 181 IATGAGIILGYSYRGKDLKEQHDQVCEREMORITPLSAFTNPCEIVDEKTVVHTS 240
DB 182 IATGAGIILGYSYRGKDLKEQHDQVCEREMORITPLSAFTNPCEIVDEKTVVHTS 241
QY 241 QTPVDPQSGSTPLMGQAGTPGA 262
DB 242 QTPVDPQSGSTPLMGQAGTPGA 263

RESULT 8
US-60-048-000-2
/ Sequence 2, Application US/60048000
/ GENERAL INFORMATION:
/ APPLICANT: EBNER, REINHARD
/ APPLICANT: MOORE, PAUL
/ APPLICANT: RUBEN, STEVE
/ TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE
/ NUMBER OF SEQUENCES: 16
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSEE: HUMAN GENOME SCIENCES, INC.
/ STREET: 9410 KEY WEST AVENUE
/ CITY: ROCKVILLE
/ STATE: MD
/ COUNTRY: USA
/ ZIP: 20850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/60/048,000
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BROOKES, ANDERS A.
/ REGISTRATION NUMBER: 36,373
/ REFERENCE/DOCKET NUMBER: PF379PP
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (301) 309-8504
/ TELEFAX: (301) 309-8512
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
```

```
/ LENGTH: 263 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-60-048-000-2

Query Match
Best Local Similarity 100.0%; Score 1388; DB 27; Length 263;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAWQAFVSNMLLAAYSGGCFWMDNGHLYREDQTSPPAGLRCLNMLDQSGIASAPV 60
DB 2 LLAWQAFVSNMLLAAYSGGCFWMDNGHLYREDQTSPPAGLRCLNMLDQSGIASAPV 61
QY 61 SGAGNHSYCRNPDDPDPGPMCYSGAGVPEKRCEDLRCPETTSQALPAFTTEIOEASE 120
DB 62 SGAGNHSYCRNPDDPDPGPMCYSGAGVPEKRCEDLRCPETTSQALPAFTTEIOEASE 121
QY 121 GGADEVGVFAPANAALPARSEAAAVOPVIGISQVRNNSKEKKDLGTIGYVLGITMVI 180
DB 122 GGADEVGVFAPANAALPARSEAAAVOPVIGISQVRNNSKEKKDLGTIGYVLGITMVI 181
QY 181 IATGAGIILGYSYRGKDLKEQHDQVCEREMORITPLSAFTNPCEIVDEKTVVHTS 240
DB 182 IATGAGIILGYSYRGKDLKEQHDQVCEREMORITPLSAFTNPCEIVDEKTVVHTS 241
QY 241 QTPVDPQSGSTPLMGQAGTPGA 262
DB 242 QTPVDPQSGSTPLMGQAGTPGA 263

RESULT 9
US-09-791-537-39562
/ Sequence 39562, Application US/09791537
/ GENERAL INFORMATION:
/ APPLICANT: Biomomix, Inc.
/ APPLICANT: Dede, Derek
/ APPLICANT: Danzer, Joseph
/ TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB.
/ FILE REFERENCE: 261/210
/ CURRENT APPLICATION NUMBER: US/09/791,537
/ CURRENT FILING DATE: 2001-02-22
/ NUMBER OF SEQ ID NOS: 153055
/ SOFTWARE: Patent version 3.0
/ SEQ ID NO: 39562
/ LENGTH: 263
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-791-537-39562

Query Match
Best Local Similarity 99.6%; Score 1385; DB 21; Length 263;
Matches 261; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAWQAFVSNMLLAAYSGGCFWMDNGHLYREDQTSPPAGLRCLNMLDQSGIASAPV 60
DB 2 LLAWQAFVSNMLLAAYSGGCFWMDNGHLYREDQTSPPAGLRCLNMLDQSGIASAPV 61
QY 61 SGAGNHSYCRNPDDPDPGPMCYSGAGVPEKRCEDLRCPETTSQALPAFTTEIOEASE 120
DB 62 SGAGNHSYCRNPDDPDPGPMCYSGAGVPEKRCEDLRCPETTSQALPAFTTEIOEASE 121
QY 121 GGADEVGVFAPANAALPARSEAAAVOPVIGISQVRNNSKEKKDLGTIGYVLGITMVI 180
DB 122 GGADEVGVFAPANAALPARSEAAAVOPVIGISQVRNNSKEKKDLGTIGYVLGITMVI 181
QY 181 IATGAGIILGYSYRGKDLKEQHDQVCEREMORITPLSAFTNPCEIVDEKTVVHTS 240
DB 182 IATGAGIILGYSYRGKDLKEQHDQVCEREMORITPLSAFTNPCEIVDEKTVVHTS 241
QY 241 QTPVDPQSGSTPLMGQAGTPGA 262
DB 242 QTPVDPQSGSTPLMGQAGTPGA 263
```

Db 242 0TRVDP0EGTTPUMQAGTPGA 263

```

RESULT 10
US-09-927-796-44
; Sequence 44, Application US/09927796
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Guiney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Masters, Scott A.
; APPLICANT: Pan, James
; APPLICANT: Pitti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Matenabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2331R1C1
; CURRENT APPLICATION NUMBER: US/09/927,796
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067411
; PRIOR FILING DATE: 1997-12-03
; PRIOR APPLICATION NUMBER: 60/069862
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/095929
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/097978
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/103396
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/108867
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/112851
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/119965
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: 60/123972
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: 60/133459
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/140650
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/140653
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698

```

```

; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 60/149395
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/151689
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/625328
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 08/710802
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 08/800699
; PRIOR FILING DATE: 1997-02-14
; PRIOR APPLICATION NUMBER: 08/828683
; PRIOR FILING DATE: 1997-03-31
; PRIOR APPLICATION NUMBER: 08/829270
; PRIOR FILING DATE: 1997-03-31
; PRIOR APPLICATION NUMBER: 08/928069
; PRIOR FILING DATE: 1997-09-11
; PRIOR APPLICATION NUMBER: 08/934494
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 09/143068
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: 09/143707
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: 09/151889
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: 09/169104
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: 09/202089
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: 09/254311
; PRIOR FILING DATE: 1999-03-03
; PRIOR APPLICATION NUMBER: 09/304003
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/380138
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/380139
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/403297
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 09/423844
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 09/511133
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 09/511631
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 09/664610
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/665350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/690169
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/690189
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/709238
; PRIOR FILING DATE: 2000-11-18
; PRIOR APPLICATION NUMBER: 09/866034
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 09/872035
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 09/884733
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: 09/886342
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: 09/866028
; PRIOR FILING DATE: 2001-08-25
; PRIOR APPLICATION NUMBER: PCT/US97/05230
; PRIOR FILING DATE: 1997-03-31
; PRIOR APPLICATION NUMBER: PCT/US98/19094
; PRIOR FILING DATE: 1998-09-14

```

```

PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: PCT/US98/21407
PRIOR FILING DATE: 1998-10-09
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: 1998-12-01
PRIOR APPLICATION NUMBER: PCT/US99/05028
PRIOR FILING DATE: 1999-03-08
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/20111
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28301
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/28634
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: PCT/US00/04341
PRIOR FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: PCT/US00/04342
PRIOR FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/06884
PRIOR FILING DATE: 2000-03-15
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: PCT/US00/13705
PRIOR FILING DATE: 2000-05-17
PRIOR APPLICATION NUMBER: PCT/US00/14941
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/20710
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: PCT/US01/17860
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: PCT/US01/19692
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: PCT/US01/21066
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: PCT/US01/21735
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 258
SEQ ID NO 44
LENGTH: 263
Query Match          99.8%; Score 1385; DB 23; Length 263;
Best Local Similarity 99.6%; Pred. No. 3,6e-128;
Matches 261; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LLAWAQAFIVSNMLLAAYAGSGCFWMDNGHLVREDQTSAPGIRJCINWMLDAOSGLASAPV 60
DB 2 LLAWAQAFIVSNMLLAAYAGSGCFWMDNGHLVREDQTSAPGIRJCINWMLDAOSGLASAPV 61
QY 61 SGAGNHSYCNPNPDDPRGWCYVSGEAGVPEKRPCEDLRCPEPTTSQALPAFTTEIOEASE 120
DB 62 SGAGNHSYCNPNPDDPRGWCYVSGEAGVPEKRPCEDLRCPEPTTSQALPAFTTEIOEASE 121
QY 121 GGAUDEVQVAPANALPARSEAAVQPVIGISQVRMNSKEKKDGTGTVLVGITMNVII 180

```

```

DB 122 GGAUDEVQVAPANALPARSEAAVQPVIGISQVRMNSKEKKDGTGTVLVGITMNVII 181
QY 181 IAIAGAGIILGYSYKRGKDLKECHDKQKVEREMORITPLSAFTNPCEIYDEKTVVHTS 240
DB 182 IAIAGAGIILGYSYKRGKDLKECHDKQKVEREMORITPLSAFTNPCEIYDEKTVVHTS 241
QY 241 QTPVDPQSGSTPLMGQAGTPGA 262
DB 242 QTPVDPQSGSTPLMGQAGTPGA 263

```

```

RESULT 11
US-10-210-951-44
Sequence 44, Application US/10210951
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Masters, Scot A.
APPLICANT: Pan, James
APPLICANT: Pitti, Robert M.
APPLICANT: Roy, Margaret Ann
APPLICANT: Smith, Victoria
APPLICANT: Stone, Donna M.
APPLICANT: Watanabe, Colin K.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
FILE REFERENCE: P2931R1C1
CURRENT APPLICATION NUMBER: US/10/210,951
CURRENT FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: 60/014699
PRIOR FILING DATE: 1996-04-01
PRIOR APPLICATION NUMBER: 60/026943
PRIOR FILING DATE: 1996-09-23
PRIOR APPLICATION NUMBER: 60/059121
PRIOR FILING DATE: 1997-07-17
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/062037
PRIOR FILING DATE: 1997-10-10
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 258
SEQ ID NO 44
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-10-210-951-44
Query Match          99.8%; Score 1385; DB 26; Length 263;
Best Local Similarity 99.6%; Pred. No. 3,6e-128;
Matches 261; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LLAWAQAFIVSNMLLAAYAGSGCFWMDNGHLVREDQTSAPGIRJCINWMLDAOSGLASAPV 60
DB 2 LLAWAQAFIVSNMLLAAYAGSGCFWMDNGHLVREDQTSAPGIRJCINWMLDAOSGLASAPV 61
QY 61 SGAGNHSYCNPNPDDPRGWCYVSGEAGVPEKRPCEDLRCPEPTTSQALPAFTTEIOEASE 120
DB 62 SGAGNHSYCNPNPDDPRGWCYVSGEAGVPEKRPCEDLRCPEPTTSQALPAFTTEIOEASE 121

```

```
QY 121 GPGADEVOVFAPANALPARSEAAVQPIGISOVRNMSKEKDLGTLGYLGTMMVII 180
DB 122 GPGADEVOVFAPANALPARSEAAVQPIGISOVRNMSKEKDLGTLGYLGTMMVII 181
QY 181 IAIAGGIIIGSYKRGKDLKEQHDQKVCEREMORITTLPLSAFTNPTCEIYDEKTVVHTS 240
DB 182 IAIAGGIIIGSYKRGKDLKEQHDQKVCEREMORITTLPLSAFTNPTCEIYDEKTVVHTS 241
QY 241 QTPVDPQEGSTPLMGQAGTPGA 262
DB 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 12
US-10-211-858-44
; Sequence 44, Application US/10211858
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Pilti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/211,858
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-858-44

Query Match 99.8%; Score 1385; DB 26; Length 263;
Best Local Similarity 99.6%; Pred. No. 3, 6e-128;
Matches 261; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAWVOAFLVSNMLLAAYGSGGCFWNGHLYREDQTSAPAGRCINWMLDAOSGLASAPV 60
DB 2 LLAWVOAFLVSNMLLAAYGSGGCFWNGHLYREDQTSAPAGRCINWMLDAOSGLASAPV 61
QY 61 SGAGNHSYCNRPEDPRGPMCYVSGEAGVPEKRPCEDLRCPETTSQALPAFTTIEOASE 120
DB 62 SGAGNHSYCNRPEDPRGPMCYVSGEAGVPEKRPCEDLRCPETTSQALPAFTTIEOASE 121
```

```
QY 121 GPGADEVOVFAPANALPARSEAAVQPIGISOVRNMSKEKDLGTLGYLGTMMVII 180
DB 122 GPGADEVOVFAPANALPARSEAAVQPIGISOVRNMSKEKDLGTLGYLGTMMVII 181
QY 181 IAIAGGIIIGSYKRGKDLKEQHDQKVCEREMORITTLPLSAFTNPTCEIYDEKTVVHTS 240
DB 182 IAIAGGIIIGSYKRGKDLKEQHDQKVCEREMORITTLPLSAFTNPTCEIYDEKTVVHTS 241
QY 241 QTPVDPQEGSTPLMGQAGTPGA 262
DB 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 13
US-10-211-884-44
; Sequence 44, Application US/10211884
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Pilti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/211,884
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-884-44

Query Match 99.8%; Score 1385; DB 26; Length 263;
Best Local Similarity 99.6%; Pred. No. 3, 6e-128;
Matches 261; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAWVOAFLVSNMLLAAYGSGGCFWNGHLYREDQTSAPAGRCINWMLDAOSGLASAPV 60
DB 2 LLAWVOAFLVSNMLLAAYGSGGCFWNGHLYREDQTSAPAGRCINWMLDAOSGLASAPV 61
QY 61 SGAGNHSYCNRPEDPRGPMCYVSGEAGVPEKRPCEDLRCPETTSQALPAFTTIEOASE 120
```

```
Db 62 SGAGNHSYCRNDEDEPRGWCYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASE 121
Qy 121 GGADEVQVFAPANALPARSEAAAVQPVIGISQVRMNSKEKDLGTLGYLGIITMVI 180
Db 122 GGADEVQVFAPANALPARSEAAAVQPVIGISQVRMNSKEKDLGTLGYLGIITMVI 181
Qy 181 IAIAGIILGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPTCEIIVDEKTVVHTS 240
Db 182 IAIAGIILGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPTCEIIVDEKTVVHTS 241
Qy 241 QTPVDPQEGSTPLMGQAGTPGA 262
Db 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 14
US-60-230-435-1580
; Sequence 1580, Application US/60230435
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: C1000768
; CURRENT APPLICATION NUMBER: US/60/230,435
; NUMBER OF SEQ ID NOS: 2991
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1580
; LENGTH: 263
; TYPE: PRP
; ORGANISM: HUMAN
US-60-230-435-1580

Query Match 99.8%; Score 1385; DB 27; Length 263;
Best Local Similarity 99.6%; Pred. No. 3.6e-128;
Matches 261; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLAWQAFVLSNMMLAEAYSGGCFWMDNGHLYREDQTSAPAGLRCLNWLDAQSGLASAPV 60
Db 2 LLAWQAFVLSNMMLAEAYSGGCFWMDNGHLYREDQTSAPAGLRCLNWLDAQSGLASAPV 61
Qy 61 SGAGNHSYCRNDEDEPRGWCYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASE 120
Db 62 SGAGNHSYCRNDEDEPRGWCYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASE 121
Qy 121 GGADEVQVFAPANALPARSEAAAVQPVIGISQVRMNSKEKDLGTLGYLGIITMVI 180
Db 122 GGADEVQVFAPANALPARSEAAAVQPVIGISQVRMNSKEKDLGTLGYLGIITMVI 181
Qy 181 IAIAGIILGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPTCEIIVDEKTVVHTS 240
Db 182 IAIAGIILGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPTCEIIVDEKTVVHTS 241
Qy 241 QTPVDPQEGSTPLMGQAGTPGA 262
Db 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 15
US-60-389-987-244
; Sequence 244, Application US/60389987
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465P2
```

```
; CURRENT APPLICATION NUMBER: US/60/389,987
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 3025
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 244
; LENGTH: 263
; TYPE: PRP
; ORGANISM: Homo sapiens
US-60-389-987-244

Query Match 99.8%; Score 1385; DB 27; Length 263;
Best Local Similarity 99.6%; Pred. No. 3.6e-128;
Matches 261; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLAWQAFVLSNMMLAEAYSGGCFWMDNGHLYREDQTSAPAGLRCLNWLDAQSGLASAPV 60
Db 2 LLAWQAFVLSNMMLAEAYSGGCFWMDNGHLYREDQTSAPAGLRCLNWLDAQSGLASAPV 61
Qy 61 SGAGNHSYCRNDEDEPRGWCYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASE 120
Db 62 SGAGNHSYCRNDEDEPRGWCYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASE 121
Qy 121 GGADEVQVFAPANALPARSEAAAVQPVIGISQVRMNSKEKDLGTLGYLGIITMVI 180
Db 122 GGADEVQVFAPANALPARSEAAAVQPVIGISQVRMNSKEKDLGTLGYLGIITMVI 181
Qy 181 IAIAGIILGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPTCEIIVDEKTVVHTS 240
Db 182 IAIAGIILGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPTCEIIVDEKTVVHTS 241
Qy 241 QTPVDPQEGSTPLMGQAGTPGA 262
Db 242 QTPVDPQEGSTPLMGQAGTPGA 263
```

Search completed: April 7, 2003, 08:49:48
Job time : 328 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 08:42:52 ; Search time 31 Seconds
(without alignments)

1278.158 Million cell updates/sec

Title: US-10-057-951-2_COPY_2_263

Perfect score: 1388
Sequence: 1 LLAAMVQAFVLSNMLAELAVG.....PVDPOEGSTPLMGQAGTPGA 262

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 679521 seqs, 151232488 residues

Total number of hits satisfying chosen parameters: 679521

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA.New:*
1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-----------------------|-------------------|
| 1 | 1379 | 99.4 | 263 | 6 US-10-218-140-6002 | Sequence 6002, Ap |
| 2 | 1115.5 | 80.4 | 264 | 6 US-10-144-779-488 | Sequence 488, Ap |
| 3 | 154.5 | 11.1 | 293 | 7 US-60-452-680-14406 | Sequence 14406, A |
| 4 | 154.5 | 11.1 | 293 | 7 US-60-453-135-8958 | Sequence 8958, Ap |
| 5 | 154.5 | 11.1 | 293 | 7 US-60-452-680-14405 | Sequence 8958, Ap |
| 6 | 154.5 | 11.1 | 516 | 7 US-60-453-135-8957 | Sequence 14405, A |
| 7 | 154.5 | 11.1 | 516 | 7 US-60-453-135-8957 | Sequence 8957, Ap |
| 8 | 154.5 | 11.1 | 527 | 5 US-09-612-314A-51 | Sequence 51, Appl |
| 9 | 154.5 | 11.1 | 527 | 5 PCT-US02-31263-26 | Sequence 26, Appl |
| 10 | 154.5 | 11.1 | 562 | 5 US-09-703-695A-4 | Sequence 4, Appl |
| 11 | 154.5 | 11.1 | 562 | 6 US-10-287-994-26 | Sequence 26, Appl |
| 12 | 154.5 | 11.1 | 562 | 7 US-60-452-680-14408 | Sequence 14408, A |
| 13 | 154.5 | 11.1 | 562 | 7 US-60-453-135-8960 | Sequence 8960, Ap |
| 14 | 154.5 | 11.1 | 562 | 7 US-60-453-135-8960 | Sequence 8960, Ap |
| 15 | 154.5 | 11.1 | 587 | 5 US-09-949-016-11501 | Sequence 11501, A |
| 16 | 154.5 | 10.6 | 655 | 1 PCT-US02-19017-28 | Sequence 28, Appl |
| 17 | 146.5 | 10.6 | 656 | 7 US-60-453-135-11516 | Sequence 11516, A |
| 18 | 146.5 | 10.6 | 656 | 7 US-60-453-135-11516 | Sequence 11516, A |
| 19 | 146.5 | 10.6 | 656 | 7 US-60-453-135-11516 | Sequence 11516, A |
| 20 | 146 | 10.5 | 322 | 1 PCT-US02-27855-21 | Sequence 21, Appl |
| 21 | 146 | 10.5 | 322 | 1 PCT-US02-27855-21 | Sequence 21, Appl |
| 22 | 146 | 10.5 | 322 | 6 US-10-233-675A-21 | Sequence 21, Appl |
| 23 | 143.5 | 10.3 | 482 | 7 US-60-452-680-14407 | Sequence 14407, A |
| 24 | 143.5 | 10.3 | 482 | 7 US-60-453-135-8959 | Sequence 8959, Ap |
| 25 | 143.5 | 10.3 | 482 | 7 US-60-453-135-8959 | Sequence 8959, Ap |
| 26 | 143 | 10.3 | 713 | 5 US-09-949-016-9983 | Sequence 9983, Ap |

| | | | | | |
|----|-------|------|-----|-----------------------|-------------------|
| 27 | 143 | 10.3 | 713 | 7 US-60-452-680-22399 | Sequence 22399, A |
| 28 | 143 | 10.3 | 713 | 7 US-60-453-135-13847 | Sequence 13847, A |
| 29 | 143 | 10.3 | 713 | 7 US-60-453-050-13847 | Sequence 13847, A |
| 30 | 142 | 10.2 | 339 | 6 US-10-401-108-3 | Sequence 3, Appl |
| 31 | 142 | 10.2 | 363 | 6 US-10-292-418-11 | Sequence 11, Appl |
| 32 | 142 | 10.2 | 391 | 1 PCT-US02-37879-7 | Sequence 7, Appl |
| 33 | 142 | 10.2 | 391 | 1 US-10-304-287-7 | Sequence 7, Appl |
| 34 | 142 | 10.2 | 394 | 1 PCT-US02-37879-8 | Sequence 8, Appl |
| 35 | 142 | 10.2 | 394 | 1 US-10-304-287-8 | Sequence 8, Appl |
| 36 | 142 | 10.2 | 453 | 5 US-09-291-200A-1 | Sequence 1, Appl |
| 37 | 142 | 10.2 | 458 | 5 US-09-946-893B-4 | Sequence 4, Appl |
| 38 | 142 | 10.2 | 569 | 5 US-09-946-893B-5 | Sequence 5, Appl |
| 39 | 142 | 10.2 | 571 | 5 US-09-946-893B-8 | Sequence 8, Appl |
| 40 | 142 | 10.2 | 576 | 5 US-09-946-893B-6 | Sequence 6, Appl |
| 41 | 142 | 10.2 | 791 | 1 PCT-US02-37879-1 | Sequence 1, Appl |
| 42 | 142 | 10.2 | 791 | 1 US-10-304-287-1 | Sequence 1, Appl |
| 43 | 142 | 10.2 | 810 | 5 US-09-291-200A-5 | Sequence 5, Appl |
| 44 | 142 | 10.2 | 810 | 5 US-09-946-893B-2 | Sequence 2, Appl |
| 45 | 141.5 | 10.2 | 688 | 1 PCT-US02-27855-18 | Sequence 18, Appl |

ALIGNMENTS

RESULT 1
US-10-218-140-6002
; Sequence 6002, Application US/10218140
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES
; FILE REFERENCE: 15966-543 CON
; CURRENT FILING DATE: 2002-08-12
; PRIOR FILING DATE: 2000-03-30
; PRIOR FILING DATE: 2000-03-30
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/127,728
; PRIOR FILING DATE: 1999-04-02
; PRIOR FILING DATE: 1999-04-02
; PRIOR APPLICATION NUMBER: 60/127,607
; NUMBER OF SEQ ID NOS: 6322
; SOFTWARE: Curanator Version 1.0
; SEQ ID NO 6002
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-218-140-6002

Query Match 99.4%; Score 1379; DB 6; Length 263;
Best Local Similarity 99.2%; Pred. No. 1.5e-110;
Matches 260; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 1 | LLAAMVQAFVLSNMLAELAVGSGGCFMDNGLHYREDOTSPPAPRLCLMWLDASGLASAPV | 60 |
| DB | 2 | LLAAMVQAFVLSNMLAELAVGSGGCFMDNGLHYREDOTSPPAPRLCLMWLDASGLASAPV | 61 |
| QY | 61 | SGAGNHSYCNPDDEDRGFCVSGAGVBEKRPCECDLRCPETTSQALPAFTTEIOEASE | 120 |
| DB | 62 | SGAGNHSYCNPDDEDRGFCVSGAGVBEKRPCECDLRCPETTSQALPAFTTEIOEASE | 121 |
| QY | 121 | GGAGDEYQVAPANALPARSEAAAVQVIGISQVRNMSKEKDLGTIGYVIGITMAYI | 180 |
| DB | 122 | GGAGDEYQVAPANALPARSEAAAVQVIGISQVRNMSKEKDLGTIGYVIGITMAYI | 181 |
| QY | 181 | IAIGAGIILGYSYKSKDKLKEQHDQVCEREMQRTPLISAFNPTCEIVDEKTVVHTS | 240 |
| DB | 182 | IAIGAGIILGYSYKSKDKLKEQHDQVCEREMQRTPLISAFNPTCEIVDEKTVVHTS | 241 |
| QY | 241 | QTPVDPOEGSTPLMGQAGTPGA 262 | |


```
; CURRENT APPLICATION NUMBER: PCT/US02/32263
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/338,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,592
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/334,233
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/334,301
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 26
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-32263-26

Query Match
Best Local Similarity 11.1%; Score 154.5; DB 1; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.5e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 24 CFWDNGHLYREDQTSAPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 74
DB 127 CYEDGISTRGWSTAESGAECTNW--NSSALAKPYSGRRPDATRLGJGNHYCRNPDR 184
DB 185 DSK-FWCYVF-KAGKYSESEFCSTPACSEGNS 213

RESULT 11
US-09-703-695A-4
; Sequence 4, Application US/09703695A
; GENERAL INFORMATION:
; APPLICANT: Xu, Yuan
; TITLE OF INVENTION: REVERSE-PHASE HPLC ASSAY FOR PLASMINOGEN ACTIVATORS
; FILE REFERENCE: P1788R1
; CURRENT APPLICATION NUMBER: US/09/703,695A
; CURRENT FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/163,607
; PRIOR FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 4
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-703-695A-4

Query Match
Best Local Similarity 11.1%; Score 154.5; DB 5; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.5e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 24 CFWDNGHLYREDQTSAPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 74
DB 127 CYEDGISTRGWSTAESGAECTNW--NSSALAKPYSGRRPDATRLGJGNHYCRNPDR 184
QY 75 DRGPMCYVSGEAGVPEKRPCEDLRCPEPTTS 105
DB 185 DSK-FWCYVF-KAGKYSESEFCSTPACSEGNS 213

RESULT 12

US-10-287-994-26
; Sequence 26, Application US/10287994
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Deftrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Bower, Caryn
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; TITLE OF INVENTION: REMODELING AND GLYCOCONJUGATION OF PEPTIDES
; FILE REFERENCE: 040853-01-5052-00
; CURRENT APPLICATION NUMBER: US/10/287,994
; CURRENT FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 26
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-994-26

Query Match
Best Local Similarity 11.1%; Score 154.5; DB 6; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.5e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 24 CFWDNGHLYREDQTSAPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 74
DB 127 CYEDGISTRGWSTAESGAECTNW--NSSALAKPYSGRRPDATRLGJGNHYCRNPDR 184
QY 75 DRGPMCYVSGEAGVPEKRPCEDLRCPEPTTS 105
DB 185 DSK-FWCYVF-KAGKYSESEFCSTPACSEGNS 213

RESULT 13
US-60-452-680-14408
; Sequence 14408, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GRUBE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CU001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14408
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-452-680-14408

Query Match
Best Local Similarity 11.1%; Score 154.5; DB 7; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.5e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 24 CFWDNGHLYREDQTSAPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 74
```

Job time : 32 secs

```
Db 127 CYEDGISTYRGWTSTAESGAECTW--NSSALAKPYSGRRPDATRLGIGNHNYCRNPDR 184
QY 75 DPRGWCYVSGAGVPEKRPCEDLRCPETTS 105
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGENS 213
```

RESULT 14

```
US-60-453-135-8960
; Sequence 8960, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CI001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8960
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-8960
```

Query Match

```
11.1%; Score 154.5; DB 7; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.5e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
```

```
QY 24 CFWDNGHLYREDQTSAPAGLRLCLNMLDAQSLASAPVS-----GAGNHSYCRNPDE 74
Db 127 CYEDGISTYRGWTSTAESGAECTW--NSSALAKPYSGRRPDATRLGIGNHNYCRNPDR 184
QY 75 DPRGWCYVSGAGVPEKRPCEDLRCPETTS 105
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGENS 213
```

RESULT 15

```
US-60-453-050-8960
; Sequence 8960, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CI001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8960
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-8960
```

Query Match

```
11.1%; Score 154.5; DB 7; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.5e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
```

```
QY 24 CFWDNGHLYREDQTSAPAGLRLCLNMLDAQSLASAPVS-----GAGNHSYCRNPDE 74
Db 127 CYEDGISTYRGWTSTAESGAECTW--NSSALAKPYSGRRPDATRLGIGNHNYCRNPDR 184
QY 75 DPRGWCYVSGAGVPEKRPCEDLRCPETTS 105
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGENS 213
```

Search completed: April 7, 2003, 08:50:27

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 08:41:52 ; Search time 15 Seconds

(without alignments)
513.920 Million cell updates/sec

Title: US-10-057-951-2_COPY_2_263

Perfect score: 1388

Sequence: 1 LLAWQAFIVSNMLAEAYG.....PVDPOGSGTPIWGQATFGA 262

Scoring table: BLOSUM62

Gapopen 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA*

1: /cgn2_6/prodata/1/1aa/5A_COMB.pep.*

2: /cgn2_6/prodata/1/1aa/5B_COMB.pep.*

3: /cgn2_6/prodata/1/1aa/6A_COMB.pep.*

4: /cgn2_6/prodata/1/1aa/6B_COMB.pep.*

5: /cgn2_6/prodata/1/1aa/6C_COMB.pep.*

6: /cgn2_6/prodata/1/1aa/6D_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 1388 | 100.0 | 263 | 4 | US-09-411-977-2 |
| 2 | 163.5 | 11.8 | 432 | 2 | US-08-811-949-63 |
| 3 | 154.5 | 11.1 | 437 | 2 | US-08-811-949-49 |
| 4 | 154.5 | 11.1 | 437 | 2 | US-08-811-949-51 |
| 5 | 154.5 | 11.1 | 437 | 2 | US-08-811-949-55 |
| 6 | 154.5 | 11.1 | 437 | 2 | US-08-811-949-57 |
| 7 | 154.5 | 11.1 | 527 | 1 | US-07-609-510B-16 |
| 8 | 154.5 | 11.1 | 527 | 1 | US-08-811-949-39 |
| 9 | 154.5 | 11.1 | 527 | 5 | PCT-US91-01025A-2 |
| 10 | 154.5 | 11.1 | 527 | 6 | 5520913-1 |
| 11 | 154.5 | 11.1 | 527 | 6 | 5520913-1 |
| 12 | 154.5 | 11.1 | 527 | 6 | 5520913-1 |
| 13 | 154.5 | 11.1 | 527 | 6 | 5520913-1 |
| 14 | 154.5 | 11.1 | 527 | 2 | US-08-811-949-43 |
| 15 | 154.5 | 11.1 | 527 | 2 | US-08-560-098A-50 |
| 16 | 154.5 | 11.1 | 527 | 2 | US-08-883-795A-38 |
| 17 | 154.5 | 11.1 | 527 | 6 | 5200340-2 |
| 18 | 154.5 | 11.1 | 527 | 6 | 5200340-2 |
| 19 | 148.5 | 10.7 | 83 | 2 | US-08-811-949-2 |
| 20 | 146.5 | 10.6 | 655 | 1 | US-08-148-910-12 |
| 21 | 146.5 | 10.6 | 655 | 1 | US-08-148-937A-12 |
| 22 | 146.5 | 10.6 | 655 | 1 | US-08-612-788-35 |
| 23 | 146.5 | 10.6 | 655 | 1 | US-08-612-788-35 |
| 24 | 146.5 | 10.6 | 655 | 1 | US-08-612-788-35 |
| 25 | 146.5 | 10.6 | 655 | 1 | US-08-612-788-35 |
| 26 | 146.5 | 10.6 | 655 | 1 | US-08-612-788-35 |
| 27 | 146.5 | 10.6 | 655 | 1 | US-08-612-788-35 |

| | | | | | | |
|----|-----|------|-----|---|------------------|--------------------|
| 28 | 142 | 10.2 | 339 | 1 | US-08-452-260-3 | Sequence 3, Appl1 |
| 29 | 142 | 10.2 | 339 | 1 | US-08-326-785-3 | Sequence 3, Appl1 |
| 30 | 142 | 10.2 | 339 | 2 | US-08-612-788-3 | Sequence 3, Appl1 |
| 31 | 142 | 10.2 | 339 | 2 | US-08-605-598B-3 | Sequence 3, Appl1 |
| 32 | 142 | 10.2 | 339 | 2 | US-08-429-743-3 | Sequence 3, Appl1 |
| 33 | 142 | 10.2 | 339 | 2 | US-08-866-735-3 | Sequence 3, Appl1 |
| 34 | 142 | 10.2 | 339 | 3 | US-09-066-028-3 | Sequence 3, Appl1 |
| 35 | 142 | 10.2 | 339 | 3 | PCT-US95-05107-3 | Sequence 3, Appl1 |
| 36 | 142 | 10.2 | 352 | 2 | US-08-612-788-40 | Sequence 40, Appl1 |
| 37 | 142 | 10.2 | 352 | 2 | US-09-066-028-40 | Sequence 40, Appl1 |
| 38 | 142 | 10.2 | 374 | 4 | US-09-377-250-3 | Sequence 40, Appl1 |
| 39 | 142 | 10.2 | 375 | 4 | US-09-377-250-2 | Sequence 42, Appl1 |
| 40 | 142 | 10.2 | 378 | 2 | US-08-612-788-42 | Sequence 42, Appl1 |
| 41 | 142 | 10.2 | 378 | 4 | US-09-066-028-42 | Sequence 42, Appl1 |
| 42 | 142 | 10.2 | 378 | 4 | US-09-206-059-1 | Sequence 1, Appl1 |
| 43 | 142 | 10.2 | 451 | 4 | US-09-377-250-1 | Sequence 1, Appl1 |
| 44 | 142 | 10.2 | 452 | 4 | US-09-377-250-4 | Sequence 4, Appl1 |
| 45 | 142 | 10.2 | 790 | 1 | US-08-469-486-54 | Sequence 54, Appl1 |

ALIGNMENTS

RESULT 1
US-09-411-977-2
Sequence 2, Application US/09411977
Patent No. 6372473
GENERAL INFORMATION:
APPLICANT: Moore, Paul A.
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
FILE REFERENCE: PF378P1
CURRENT APPLICATION NUMBER: US/09/411,977
EARLIER FILING DATE: 1999-10-04
EARLIER APPLICATION NUMBER: 09/084,491
EARLIER FILING DATE: 1998-05-27
EARLIER APPLICATION NUMBER: 60/048,000
EARLIER FILING DATE: 1997-05-28
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-09-411-977-2

Query Match 100.0%; Score 1388; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 5.4e-139;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 1 | LLAWQAFIVSNMLAEAYGSGCFVNDGHIYRDOTSPAPGTCIMWLDNOSGLASAV | 60 |
| DB | 2 | LLAWQAFIVSNMLAEAYGSGCFVNDGHIYRDOTSPAPGTCIMWLDNOSGLASAV | 61 |
| QY | 61 | SGAGNSYCRNDEDPGKPCVYSGAGVDEKPCEDLRCPETTSQALPAFTTEIQASE | 120 |
| DB | 62 | SGAGNSYCRNDEDPGKPCVYSGAGVDEKPCEDLRCPETTSQALPAFTTEIQASE | 121 |
| QY | 121 | GPADDEVQVFAFANALPARSEAAAVPVIGISQRYRMSKEKKDGLTGLTMMVIT | 180 |
| DB | 122 | GPADDEVQVFAFANALPARSEAAAVPVIGISQRYRMSKEKKDGLTGLTMMVIT | 181 |
| QY | 181 | ITAGAGIIGYGYKKGKDKCEKQKCEKQKCEKQKCEKQKCEKQKCEKQKCEKQK | 240 |
| DB | 182 | ITAGAGIIGYGYKKGKDKCEKQKCEKQKCEKQKCEKQKCEKQKCEKQKCEKQK | 241 |
| QY | 241 | QTPVDPGSGTPIWGQATFGA 262 | |
| DB | 242 | QTPVDPGSGTPIWGQATFGA 263 | |

RESULT 2

NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-51

Query Match 11.1%; Score 154.5; DB 2; Length 437;
Best Local Similarity 39.6%; Pred. No. 5.7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

OY 24 CFWDGHLVREDOTSPAPGLRCLNMLDAOSGLASAPVS-----GAGNHSYCRNPDE 74
DB 2 CYEDGISTYRGWTSTAESGAECTNW--NSSALAKPYSGRRPDARLGLGNHNYCRNPDR 59
OY 75 DRRGPMCVSGEAGVPEKRPCEDLRCPEPTS 105
DB 60 DSK-PWCTVF-KAGKYSSEFCSTPACSEGN 88

RESULT 5
US-08-811-949-55
Sequence 55, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIMA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, YOUTI
APPLICANT: KOBAYASHI, MASARAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-55

Query Match 11.1%; Score 154.5; DB 2; Length 437;

Best Local Similarity 39.6%; Pred. No. 5.7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
OY 24 CFWDGHLVREDOTSPAPGLRCLNMLDAOSGLASAPVS-----GAGNHSYCRNPDE 74
DB 2 CYEDGISTYRGWTSTAESGAECTNW--NSSALAKPYSGRRPDARLGLGNHNYCRNPDR 59
OY 75 DRRGPMCVSGEAGVPEKRPCEDLRCPEPTS 105
DB 60 DSK-PWCTVF-KAGKYSSEFCSTPACSEGN 88

RESULT 6
US-08-811-949-57
Sequence 57, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIMA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, YOUTI
APPLICANT: KOBAYASHI, MASARAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-57

Query Match 11.1%; Score 154.5; DB 2; Length 437;
Best Local Similarity 39.6%; Pred. No. 5.7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

OY 24 CFWDGHLVREDOTSPAPGLRCLNMLDAOSGLASAPVS-----GAGNHSYCRNPDE 74
DB 2 CYEDGISTYRGWTSTAESGAECTNW--NSSALAKPYSGRRPDARLGLGNHNYCRNPDR 59
OY 75 DRRGPMCVSGEAGVPEKRPCEDLRCPEPTS 105
DB 60 DSK-PWCTVF-KAGKYSSEFCSTPACSEGN 88

RESULT 7
US-07-609-510B-16
Sequence 16, Application US/076095108

```
/ Patent No. 5326700
/ GENERAL INFORMATION:
/ APPLICANT: Berg et al.
/ TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue P
/ NUMBER OF SEQUENCES: 16
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Eli Lilly and Company
/ STREET: Lilly Corporate Center
/ CITY: Indianapolis
/ STATE: IN.
/ COUNTRY: U.S.A.
/ ZIP: 46285
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB storage
/ OPERATING SYSTEM: Macintosh
/ SOFTWARE: Microsoft Word
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/609,510B
/ FILING DATE: 19901106
/ CLASSIFICATION: 435
/ INFORMATION FOR SEQ ID NO: 16:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 527 amino acids
/ TYPE: AMINO ACID
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-07-609-510B-16

Query Match 11.1%; Score 154.5; DB 1; Length 527;
Best Local Similarity 39.6%; Pred. No. 7.6e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

Qy 24 CFMDNGHLREDDTSPAPGRCLNWLDAQSLAPVS-----GAGNHSYCRNDE 74
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 92 CYEDGISTYRGTWSTAESGAECTNW--NSSALAKPYSGRRPDARLGLGNHNYCRNPD 149
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 75 DPRGPMCVSGEAGVPEKRPCELDRCPEPTTS 105
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 150 DSK-PMCYVF-KAGKYSSEFCSTPACSEGNS 178
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
US-08-811-949-39
/ Sequence 39, Application US/08811949
/ Patent No. 5840533
/ GENERAL INFORMATION:
/ APPLICANT: NIMA, MINEO
/ APPLICANT: SAITO, YOSHIMASA
/ APPLICANT: SASAKI, HITOSHI
/ APPLICANT: HAYASHI, MASAKO
/ APPLICANT: NOTANI, JOUJI
/ APPLICANT: KOBAYASHI, MASAKAZU
/ TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
/ NUMBER OF SEQUENCES: 67
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: OBLON, SPIVAK, MCCLEBLAND, MAIER & NEUSTADT,
/ ADDRESS: P.C.
/ STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
/ CITY: ARLINGTON
/ STATE: VA
/ COUNTRY: USA
/ ZIP: 22202
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/811,949
/ FILING DATE: 05-MAR-1997
/ CLASSIFICATION: 435
```

```
/ ATTORNEY/AGENT INFORMATION:
/ NAME: OBLON, NORMAN F.
/ REGISTRATION NUMBER: 24,618
/ REFERENCE/DOCKET NUMBER: 18-966-0
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 703-413-3000
/ TELEFAX: 703-413-2220
/ INFORMATION FOR SEQ ID NO: 39:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 527 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-811-949-39

Query Match 11.1%; Score 154.5; DB 2; Length 527;
Best Local Similarity 39.6%; Pred. No. 7.6e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

Qy 24 CFMDNGHLREDDTSPAPGRCLNWLDAQSLAPVS-----GAGNHSYCRNDE 74
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 92 CYEDGISTYRGTWSTAESGAECTNW--NSSALAKPYSGRRPDARLGLGNHNYCRNPD 149
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 75 DPRGPMCVSGEAGVPEKRPCELDRCPEPTTS 105
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 150 DSK-PMCYVF-KAGKYSSEFCSTPACSEGNS 178
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
PCT-US91-01025A-2
/ Sequence 2, Application PC/TUS9101025A
/ GENERAL INFORMATION:
/ APPLICANT: Genentech, Inc.
/ TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin
/ TITLE OF INVENTION: Specific Properties
/ NUMBER OF SEQUENCES: 2
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Genentech, Inc.
/ STREET: 460 Point San Bruno Blvd
/ CITY: South San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94080
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: patin (Genentech)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US91/01025A
/ FILING DATE: 19910214
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/486,657
/ FILING DATE: 1 March 1990
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Haasak, Janet E.
/ REGISTRATION NUMBER: 28,616
/ REFERENCE/DOCKET NUMBER: 454P2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415/266-1896
/ TELEFAX: 415/952-9881
/ TELE: 910/371-7168
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 527 amino acids
/ TYPE: AMINO ACID
/ TOPOLOGY: linear
/ PCT-US91-01025A-2

Query Match 11.1%; Score 154.5; DB 5; Length 527;
Best Local Similarity 39.6%; Pred. No. 7.6e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
```


Qy 24 CFWDNGHLTYREDQTSAPAGLCRLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 74
Db 92 CYEDQGISYRGTWSTAESGAECTNW--NSSALAQKPYSGRRPDARLGLGNHNYCRNPDR 149
Qy 75 DPRGPMCVYSGEAGVPEKRPCEDLRCPEPTS 105
Db 150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 178

RESULT 10
5185259-8
Patent No. 5185259
APPLICANT: GOEBDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
VEHAR, GORDON A.
TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
ACTIVATOR
NUMBER OF SEQUENCES: 15
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/489,855
FILING DATE: 02-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 12,694
FILING DATE: 09-FEB-1987
APPLICATION NUMBER: 483,052
FILING DATE: 07-APR-1983
APPLICATION NUMBER: 398,003
FILING DATE: 14-JUL-1982
APPLICATION NUMBER: 374,860
FILING DATE: 05-MAY-1982
SEQ ID NO: 8
LENGTH: 527

Query Match 11.1%; Score 154.5; DB 6; Length 527;
Best Local Similarity 39.6%; Pred. No. 7.6e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
Qy 24 CFWDNGHLTYREDQTSAPAGLCRLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 74
Db 92 CYEDQGISYRGTWSTAESGAECTNW--NSSALAQKPYSGRRPDARLGLGNHNYCRNPDR 149
Qy 75 DPRGPMCVYSGEAGVPEKRPCEDLRCPEPTS 105
Db 150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 178

RESULT 11
5520913-1
Patent No. 5520913
APPLICANT: ANDERSON, STEPHEN; BENNETT, WILLIAM F.; BOSTEIN,
DAVID; HIGGINS, DEBORAH L.; PRONI, NICHOLAS F.; ZOLLER, MARK J.
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR HAVING
ZYMOGENIC PROPERTIES
NUMBER OF SEQUENCES: 35
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/88,451
FILING DATE: 06-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 770,510
FILING DATE: 03-OCT-1991
APPLICATION NUMBER: 384,608
FILING DATE: 24-JUL-1989
APPLICATION NUMBER: 240,856
FILING DATE: 02-SEP-1988
SEQ ID NO: 1
LENGTH: 527

Query Match 11.1%; Score 154.5; DB 6; Length 527;
Best Local Similarity 39.6%; Pred. No. 7.6e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

Qy 24 CFWDNGHLTYREDQTSAPAGLCRLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 74
Db 92 CYEDQGISYRGTWSTAESGAECTNW--NSSALAQKPYSGRRPDARLGLGNHNYCRNPDR 149
Qy 75 DPRGPMCVYSGEAGVPEKRPCEDLRCPEPTS 105
Db 150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 178

RESULT 12
5200340-6
Patent No. 5200340
APPLICANT: FOSTER, DONALD C.; MULVHILL, EILEEN R.; O'HARA,
PATRICK J.; PINDEL, KURT; YOSHITAKE, SHINJI
TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
ACTIVATORS
NUMBER OF SEQUENCES: 34
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/53,412
FILING DATE: 22-MAY-1987
SEQ ID NO: 6
LENGTH: 546

Query Match 11.1%; Score 154.5; DB 6; Length 546;
Best Local Similarity 39.6%; Pred. No. 8e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
Qy 24 CFWDNGHLTYREDQTSAPAGLCRLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 74
Db 127 CYEDQGISYRGTWSTAESGAECTNW--NSSALAQKPYSGRRPDARLGLGNHNYCRNPDR 184
Qy 75 DPRGPMCVYSGEAGVPEKRPCEDLRCPEPTS 105
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213

RESULT 13
US-08-811-949-43
Sequence 43, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NITWA, MINO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C., JEFFERSON DAVIS HIGHWAY, SUITE 400
STREET: 1755 S.
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-43

Query Match 11.1%; Score 154.5; DB 2; Length 562;
Best Local Similarity 39.6%; Pred. No. 8,4e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 24 CFWDNGHLTYREDQTPAPGLRCINWLDASGLASAPVS-----GAGNHSYCRNPDE 74
DB 127 CYEDQGISYRGSTWSTAESGAECTNW--NSSALAKPYSGRRPDARLGLGNHNYCRNPDR 184

QY 75 DPGPWCYVSGEAGVPEKRPCEDLRCPEPTTS 105
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213

RESULT 14
US-08-560-098A-50
Sequence 50, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WENDET, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKewen, Edwards & Ienahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-50

Query Match 11.1%; Score 154.5; DB 2; Length 562;
Best Local Similarity 39.6%; Pred. No. 8,4e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
QY 24 CFWDNGHLTYREDQTPAPGLRCINWLDASGLASAPVS-----GAGNHSYCRNPDE 74

DB 127 CYEDQGISYRGSTWSTAESGAECTNW--NSSALAKPYSGRRPDARLGLGNHNYCRNPDR 184
QY 75 DPGPWCYVSGEAGVPEKRPCEDLRCPEPTTS 105
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213

RESULT 15
US-08-883-795A-38
Sequence 38, Application US/08883795A
Patent No. 5985607
GENERAL INFORMATION:
APPLICANT: Delcove, Genevieve
APPLICANT: Awang, Gregor
TITLE OF INVENTION: Recombinant DNA Molecules and Expression
TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3T2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,795A
FILING DATE: 27-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Micheline
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 7841-062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein (tpa)
US-08-883-795A-38

Query Match 11.1%; Score 154.5; DB 2; Length 562;
Best Local Similarity 39.6%; Pred. No. 8,4e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
QY 24 CFWDNGHLTYREDQTPAPGLRCINWLDASGLASAPVS-----GAGNHSYCRNPDE 74
DB 127 CYEDQGISYRGSTWSTAESGAECTNW--NSSALAKPYSGRRPDARLGLGNHNYCRNPDR 184
QY 75 DPGPWCYVSGEAGVPEKRPCEDLRCPEPTTS 105
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213

Search completed: April 7, 2003, 08:44:14
Job time : 17 secs

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:25 / Search time 19.6216 Seconds
(without alignments)
1185.658 Million cell updates/sec

Title: US-10-057-951-2_COPY_22_263

Perfect score: 1289

Sequence: 1 SGCGFMDNGHLYREDQTSPTA.....PVDPOEGSTPLMCAQAGTPGA 242

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 154.5 | 12.0 | 562 | 1 UKHUT | t-plasminogen acti |
| 2 | 154 | 11.9 | 559 | 1 A35029 | t-plasminogen acti |
| 3 | 147.5 | 11.4 | 559 | 1 A29941 | t-plasminogen acti |
| 4 | 146.5 | 11.4 | 655 | 1 A46688 | hepatocyte growth |
| 5 | 145.5 | 11.3 | 291 | 2 A38098 | t-plasminogen acti |
| 6 | 142 | 11.0 | 810 | 1 PLHU | plasmin (EC 3.4.21 |
| 7 | 138 | 10.7 | 169 | 2 A40522 | plasmin (EC 3.4.21 |
| 8 | 138 | 10.7 | 433 | 1 UKBAY | u-plasminogen acti |
| 9 | 135.5 | 10.5 | 442 | 1 UKPG | u-plasminogen acti |
| 10 | 134 | 10.4 | 431 | 2 US0599 | t-plasminogen acti |
| 11 | 134 | 10.4 | 477 | 2 A34369 | t-plasminogen acti |
| 12 | 134 | 10.4 | 477 | 2 US0598 | t-plasminogen acti |
| 13 | 133.5 | 10.4 | 716 | 1 JC5061 | macrophage-stimula |
| 14 | 132 | 10.2 | 431 | 1 UKHU | u-plasminogen acti |
| 15 | 130.5 | 10.1 | 716 | 1 A40332 | macrophage-stimula |
| 16 | 129 | 10.0 | 810 | 2 A46260 | plasmin (EC 3.4.21 |
| 17 | 127 | 9.9 | 394 | 2 JS0600 | t-plasminogen acti |
| 18 | 127 | 9.9 | 433 | 1 A40560 | u-plasminogen acti |
| 19 | 126.5 | 9.8 | 728 | 1 A60185 | hepatocyte growth |
| 20 | 124.5 | 9.7 | 810 | 2 B30848 | plasmin (EC 3.4.21 |
| 21 | 124 | 9.6 | 622 | 1 TBHU | thrombin (EC 3.4.2 |
| 22 | 124 | 9.6 | 728 | 1 JH0579 | hepatocyte growth |
| 23 | 124 | 9.6 | 812 | 1 PLBO | plasmin (EC 3.4.21 |
| 24 | 123.5 | 9.6 | 434 | 1 A35005 | u-plasminogen acti |
| 25 | 123 | 9.5 | 593 | 2 S45281 | coagulation factor |
| 26 | 123.5 | 9.5 | 4548 | 1 S00657 | apoptotone (a) (EC |
| 27 | 122 | 9.5 | 603 | 2 S28941 | coagulation factor |
| 28 | 121.5 | 9.4 | 728 | 1 A35644 | hepatocyte growth |
| 29 | 120 | 9.3 | 790 | 1 PLPG | plasmin (EC 3.4.21 |

| | | | | | |
|----|-------|-----|------|----------|--------------------|
| 30 | 119.5 | 9.3 | 477 | 2 JS0597 | t-plasminogen acti |
| 31 | 119.5 | 9.3 | 560 | 1 JC4795 | plasma hyaluronan |
| 32 | 119 | 9.2 | 433 | 1 UKMS | u-plasminogen acti |
| 33 | 118.5 | 9.2 | 615 | 1 KFHU12 | coagulation factor |
| 34 | 118 | 9.2 | 432 | 1 S18932 | u-plasminogen acti |
| 35 | 118 | 9.2 | 710 | 1 JS1283 | hepatocyte growth |
| 36 | 117 | 9.1 | 625 | 1 TBBO | thrombin (EC 3.4.2 |
| 37 | 116 | 9.0 | 1420 | 2 A32869 | apolipoprotein(a) |
| 38 | 115.5 | 9.0 | 711 | 1 A47136 | macrophage-stimula |
| 39 | 114.5 | 8.9 | 812 | 1 PLMS | plasmin (EC 3.4.21 |
| 40 | 114 | 8.8 | 458 | 2 A35827 | thrombin (EC 3.4.2 |
| 41 | 112 | 8.7 | 455 | 2 A61545 | plasmin (EC 3.4.21 |
| 42 | 110.5 | 8.6 | 685 | 1 A46289 | neurotrophic recep |
| 43 | 110 | 8.5 | 558 | 2 JC5878 | plasma hyaluronan |
| 44 | 108.5 | 8.4 | 123 | 2 C61545 | plasmin (EC 3.4.21 |
| 45 | 107 | 8.3 | 617 | 2 S10511 | thrombin (EC 3.4.2 |

ALIGNMENTS

RESULT 1

UKHUT
t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human
N:Alternate names: t-PA; tissue plasminogen activator
C:Species: Homo sapiens (man)
C>Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text change 08-Dec-2000
C/Accession: A94004; A23529; J05062; A93293; S02125; A91343; A93951; A91322; A54645; I60
R:NY, T.; Bligh, F.; Lund, B.
Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
A>Title: The structure of the human tissue-type plasminogen activator gene: correlation
A/Reference number: A94004; MUID:84298137; PMID:6089198
A/Accession: A94004
A/Molecule type: DNA
A/Residues: 1-562 <NT>
A/Cross-references: GB:L00141
A/Note: The codon given for residue 93 (ACC) is inconsistent with the authors' translati.
R:Friener Degen, S.J.; Rajput, B.; Retch, E.
J. Biol. Chem. 261, 6972-6985, 1986
A>Title: The human tissue plasminogen activator gene.
A/Reference number: A23529; MUID:86196143; PMID:3009482
A/Accession: A23529
A/Molecule type: DNA
A/Residues: 1-562 <DE>
A/Cross-references: GB:K03021; NID:G339817; PIDN:AAA98809.1; PID:G339818
R:Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuda, S.; Higashio, K.
Agric. Biol. Chem. 55, 1225-1232, 1991
A>Title: Purification and characterization of tissue plasminogen activator secreted by h
A/Reference number: J05062; MUID:91291340; PMID:1368681
A/Accession: J05062
A/Molecule type: mRNA
A/Residues: 31-562 <IT>
A/Cross-references: DDBJ:D01096; NID:G220128; PIDN:BA00881.1; PID:G441174
A/Experimental source: embryonic lung fibroblast IM-90 cells
A/Note: Part of this sequence, including the amino end of the mature protein, was confir
R:Penica, P.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vehar, G.A.; Ward, C.A.; Bennett
Nature 301, 214-221, 1983
A>Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche
A/Reference number: A93293; MUID:83115262; PMID:6337343
A/Accession: A93293
A/Molecule type: mRNA
A/Residues: 1-562 <PEN>
A/Cross-references: GB:L00141
A/Experimental source: melanoma cells
R:Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.
Nucleic Acids Res. 16, 5695, 1988
A>Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe
A/Reference number: S02125; MUID:88262579; PMID:3133640
A/Accession: S02125
A/Status: translation not shown
A/Molecule type: mRNA
A/Residues: 1-562 <SAS>
A/Cross-references: EMBL:X07393; NID:G37243; PIDN:CMA0302.1; PID:G37244

A:Experimental source: fetal lung cells
 R:Kakitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Ma
 FBS Lett. 189, 145-149, 1985
 A:Title: Expression in *Baculovirus* coli of finger-domain lacking tissue-type plasminogen
 A:Reference number: A91343; MUID:85285620; PMID:3896853
 A:Accession: A91343
 A:Molecule type: mRNA
 A:Residues: 1-38, 'G', 86-433, 'E', 435-562 <KAG>
 A:Experimental source: Detroit 562 cells; ATCC 138
 R:Edlund, T.; Ny, T.; Randy, M.; Hedden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
 Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
 A:Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activator
 A:Reference number: A93951; MUID:83169656; PMID:6572897
 A:Accession: A93951
 A:Molecule type: mRNA
 A:Residues: 251-358 <EDL>
 A:Experimental source: melanoma cells
 R:Pohl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jorvall, H.
 Biochemistry 23, 3701-3707, 1984
 A:Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived am
 differences.
 A:Reference number: A90486; MUID:85000468; PMID:6433976
 A:Contents: annotation; melanoma cells; partial sequence of residues 35-562, active and
 R:Pohl, G.; Kaplan, L.; Binarsson, M.; Wallen, P.; Jorvall, H.
 FBS Lett. 168, 29-32, 1984
 A:Title: Differences between uterine and melanoma forms of tissue plasminogen activator.
 A:Reference number: A91322; MUID:84158556; PMID:6538514
 A:Accession: A91322
 A:Molecule type: protein
 A:Residues: 33-45;311-320 <POH>
 A:Experimental source: uterus
 A:Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln
 R:van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.
 J. Biol. Chem. 261, 14214-14218, 1986
 A:Reference number: A37567; MUID:87033611; PMID:3021732
 A:Contents: annotation; fibrin binding site
 R:Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Enger
 EMBO J. 5, 3525-3530, 1986
 A:Title: Involvement of finger domain and kringles 2 domain of tissue-type plasminogen ac
 A:Reference number: A37568; MUID:87161761; PMID:3030730
 A:Contents: annotation; fibrin binding site
 R:Dodd, I.; Nunn, B.; Robinson, J.H.
 Thromb. Haemost. 59, 523-528, 1988
 A:Title: Isolation, identification and pharmacokinetic properties of human tissue-type p
 A:Reference number: A60902; MUID:89044681; PMID:3142086
 A:Contents: annotation; novel forms of expressed recombinant t-PA
 R:Hartley, T.J.R.; Patel, T.; Watson, F.A.O.; Little, S.; Emtage, J.S.; Opdenakker, G.;
 Mol. Biol. Med. 3, 279-292, 1986
 A:Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its exp
 A:Reference number: A54645; MUID:86284200; PMID:3090401
 A:Accession: A54645
 A:Molecule type: mRNA
 A:Residues: 1-562 <HAR>
 A:Cross-references: GB:M15518; NID:9190031; PIDN:AAA60111.1; PID:9190032
 A:Note: Parts of this sequence were confirmed by peptide sequencing
 R:Reddy, V.B.; Garramone, A.J.; Saak, H.; Wei, C.
 DNA 6, 461-472, 1987
 A:Title: Expression of human uterine tissue-type plasminogen activator in mouse cells us
 A:Reference number: I60110; MUID:88054470; PMID:2824147
 A:Accession: I60110
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-562 <RES>
 A:Cross-references: GB:M18182; NID:9340176; PIDN:AAA36800.1; PID:9340177
 R:Flaherty, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.
 J. Biol. Chem. 260, 11223-11230, 1985
 A:Title: Isolation and characterization of the human tissue-type plasminogen activator B
 A:Reference number: I55232; MUID:85289338; PMID:3161893
 A:Accession: I55232
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-36 <RE2>
 A:Cross-references: GB:M11890; NID:9339837; PIDN:AAA61213.1; PID:9339839

C:Comment: Cleavage by plasmin or trypsin produces two chains held together by a single
 C:Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond. It
 C:Comment: t-PA binds chain A of fibrin by kringles 2 and the fibronectin type I repeat.
 C:Genetics:
 A:Gene: GDB:PLAT
 A:Cross-references: GDB:119496; OMIM:173370
 A:Map position: 8p12-8p12
 A:Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3; 510
 C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringles; plasma; serine proteinase
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-33/Domain: signal sequence #status predicted <PRO>
 F:33-562/Product: t-plasminogen activator #status experimental <MNT>
 F:33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
 F:41-78/Domain: fibronectin type I repeat homology <IFI>
 F:86-119/Domain: EGF homology <EGF>
 F:127-208/Domain: kringles homology <KRI>
 F:215-296/Domain: kringles homology <KR2>
 F:311-562/Product: t-plasminogen activator chain B #status experimental <BCH>
 F:41-71,68-78,86-97,91-108,110-119,127-208,148-190,179-203,215-226,236-278,267-291,299-4
 F:152,483/Binding site: carbohydrate (asn) (covalent) #status experimental
 F:219/Binding site: carbohydrate (asn) (covalent) (partial) #status experimental
 F:310-311/Cleavage site: Arg-116 (plasmin, trypsin) #status experimental
 F:357,406/Active site: His, Asp #status predicted
 F:513/Active site: Ser #status experimental
 Query Match 12.0%; Score 154.5; DB 1; Length 562;
 Best Local Similarity 39.6%; Pred. No. 2.5e-05;
 Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
 Db 4 CFMDGNHLVREDTSPAPGRLCNWLDQAQSGLASAPVS-----GAGNYSYCNPNDE 54
 127 CYEDQGISRTGWTSTVSGAECTNM--NSSALAKRPSRRRDARLGLGNHYCRNPR 184
 QY 55 DPRGPMCVYSGEAGVPEKRPCEDLRCPEETS 85
 Db 185 DSK-PMCVYF-KAGKYSSEKCPSPACSEGNS 213
 RESULT 2
 A35029
 C-Plasminogen activator (BC 3.4.21.68) precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A35029; A31597
 R:Feng, P.; Ohlsson, M.; Ny, T.
 J. Biol. Chem. 265, 2022-2027, 1990
 A:Title: The structure of the TARA-less rat tissue-type plasminogen activator gene. Spec
 A:Reference number: A35029; MUID:90130448; PMID:2105315
 A:Accession: A35029
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-559 <FEN>
 A:Cross-references: GB:M31197; NID:9207429; PIDN:AAA42261.1; PID:9207431; GB:J05226
 R:Wu, T.; Leonardson, G.; Heuvel, A.J.W.
 DNA 7, 671-677, 1988
 A:Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activator
 A:Reference number: A31597; MUID:89170114; PMID:3148445
 A:Accession: A31597
 A:Molecule type: mRNA
 A:Residues: 1-379, 'K', 381-559 <NYT>
 A:Cross-references: GB:M23697; NID:9530159; PIDN:AAA41812.1; PID:9530160
 C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringles; serine proteinase
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-29/Domain: propeptide #status predicted <PRO>
 F:30-559/Product: t-plasminogen activator #status predicted <MNT>
 F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
 F:38-75/Domain: fibronectin type I repeat homology <IFI>
 F:83-116/Domain: EGF homology <EGF>
 F:124-205/Domain: kringles homology <KRI>
 F:213-294/Domain: kringles homology <KR2>

F:309-555/Domain: t-plasminogen activator chain B #status predicted <BCH>
F:309-555/Domain: t-rypsin homology <TRY>
F:308-66-75, 83-94, 88-105, 107-116, 124-205, 145-187, 176-200, 213-294, 234-276, 265-289, 297-4
F:149, 481/Binding site: carbohydrate (asn) (covalent) #status predicted
F:308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
F:355, 404, 510/Active site: His, Asp, Ser #status predicted

Query Match 11.9%; Score 154; DB 1; Length 559;
Best Local Similarity 32.4%; Pred. No. 2,86-05;
Matches 45; Conservative 14; Mismatches 54; Indels 26; Gaps 6;

QY 4 CPWDGHLHYREDQSPAPGRLCLNTMDAOSGLASAPVS-----GAGNHSYCRNPDE 54
DB 124 CPEGGITRYGTWSTASGAECLINW--NSSALSKPYARRPNKIKGLGNHYCRNPDR 181
QY 55 DPRGMCVYSGEAGVPEKRPCEDLRCPE-----TTSQALPAPFTTIEQDASBFG 103
DB 182 DPK-PWCYVP-KAKGKTYTFCSTPACPKPTEDCVYGVKGYTRGTHSFTT--SKASCLFW 237
QY 104 ADEVQVPAFANALPARSEA 122
DB 238 NSMILIGKTYTAMRANQA 256

RESULT 3

C:Plasminogen activator (EC 3.4.21.68) precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A29941; S48205; S48207; S48206
R:RICKLES, R.J.; DARROW, A.L.; STRICKLAND, S.
J. Biol. Chem. 263, 1563-1569, 1988
A:Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA
A:Reference number: A29941; MID:88087303; PMID:2826484
A:Accession: A29941

A:Molecule type: mRNA

A:Residues: 1-559 <RUC>

A:Cross-references: GB:J03520; NID:G202109; PIDD:AAA0470.1; PID:G202110
R:LiJuen, H.R.; van Hoet, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994

A:Title: Characterization of the murine plasma fibrinolytic system.
A:Reference number: S48202; MID:95010076; PMID:7523120

A:Accession: S48205

A:Molecule type: protein

A:Residues: 33-37, 'X', 39-40 <LIJ>

A:Accession: S48207

A:Molecule type: protein

A:Residues: 309-316 <LIJ>

A:Accession: S48206

A:Molecule type: protein

A:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringlike; serine proteinase
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-29/Domain: propeptide #status predicted <PRO>

F:30-559/Product: t-plasminogen activator #status predicted <MAT>
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>

F:38-75/Domain: fibronectin type I repeat homology <FPI>
F:83-116/Domain: EGF homology <EGF>

F:124-205/Domain: kringlike homology <KRI>
F:213-294/Domain: kringlike homology <KR2>

F:309-555/Domain: t-plasminogen activator chain B #status predicted <BCH>
F:309-555/Domain: t-rypsin homology <TRY>

F:38-66-75, 83-94, 88-105, 107-116, 124-205, 145-187, 176-200, 213-294, 234-276, 265-289, 297-4
F:149, 481/Binding site: carbohydrate (asn) (covalent) #status predicted
F:308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted

F:355, 404, 510/Active site: His, Asp, Ser #status predicted

Query Match 11.4%; Score 147.5; DB 1; Length 559;
Best Local Similarity 37.0%; Pred. No. 9,96-05;
Matches 34; Conservative 11; Mismatches 34; Indels 13; Gaps 4;

QY 4 CPWDGHLHYREDQSPAPGRLCLNTMDAOSGLASAPVS-----GAGNHSYCRNPDE 54

DB 124 CPEGGITRYGTWSTASGAECLINW--NSSVLSTKPYARRPNKIKGLGNHYCRNPDR 181
QY 55 DPRGMCVYSGEAGVPEKRPCEDLRCPE-----TTSQALPAPFTTIEQDASBFG 103
DB 182 DPK-PWCYVP-KAKGKTYTFCSTPACPKPTEDCVYGVKGYTRGTHSFTT--SKASCLFW 237

Query Match 11.4%; Score 146.5; DB 1; Length 559;
Best Local Similarity 36.9%; Pred. No. 0,00014;
Matches 41; Conservative 10; Mismatches 41; Indels 19; Gaps 6;

QY 4 CPWDGHLHYREDQSPAPGRLCLNTMDAOSGLASAPVS-----GAGNHSYCRNPDE 55
DB 286 CPEGGITRYGTWSTASGAECLINW--NSSALSKPYARRPNKIKGLGNHYCRNPDR 344
QY 56 DPRGMCVYSGEAGVPEKRPCEDLRCPE-----TTSQALPAPFTTIEQDASBFG 101
DB 345 DPK-PWCYVP-KAKGKTYTFCSTPACPKPTEDCVYGVKGYTRGTHSFTT--SKASCLFW 237

RESULT 5

C:Plasminogen activator precursor, inactive endothelial splice form - human
C:Species: Homo sapiens (man)
C:Date: 21-Sep-1993 #sequence_revision 25-Aug-1995 #text_change 08-Dec-2000
C:Accession: A46688
R:MiYazawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.
J. Biol. Chem. 268, 10024-10028, 1993
A:Title: Molecular cloning and sequence analysis of the cDNA for a human serine protease
A:Reference number: A46688; MID:93252878; PMID:7683665
A:Accession: A46688

A:Molecule type: mRNA

A:Residues: 1-655 <MIY>

A:Cross-references: DDBJ:D14012; NID:G219680; PIDD:BA03113.1; PID:G219681
A:Experimental source: liver (mRNA); serum (protein)
A:Note: sequence extracted from NCBI backbone (NCBI:131227, NCBI:131228)

A:Genetics: parts of the sequence, including the amino ends of the heavy and light chains, c
A:Gene: GDB:HGFA; HGFA; HGFA
A:Cross-references: GDB:9954514
A:Map position: 4p16-4p16

A:Function: activates hepatocyte growth factor by specific proteolytic cleavage
A:Pathway: tissue repair and regeneration
C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C:Keywords: glycoprotein; hydrolase; kringlike; liver; plasma; serine proteinase
F:1-34/Domain: signal sequence #status predicted <SIG>

F:104-148/Domain: fibronectin type II repeat homology <IF2>
F:164-197/Domain: EGF homology <EG1>

F:202-237/Domain: fibronectin type I repeat homology <FPI>
F:245-278/Domain: EGF homology <EG2>

F:386-367/Domain: kringlike homology <KR2>
F:373-407/Product: hepatocyte growth factor activator light chain #status experimental <
F:408-655/Product: hepatocyte growth factor activator heavy chain #status experimental <
F:408-641/Domain: trypsin homology <TRY>

F:40, 48, 290, 468, 492, 546/Binding site: carbohydrate (asn) (covalent) #status predicted
F:154-175, 169-186, 188-197, 202-230, 228-237, 245-256, 250-267, 269-278, 286-367, 307-349, 338-36
F:447, 497, 598/Active site: His, Asp, Ser #status predicted

Query Match 11.4%; Score 146.5; DB 1; Length 655;
Best Local Similarity 36.9%; Pred. No. 0,00014;
Matches 41; Conservative 10; Mismatches 41; Indels 19; Gaps 6;

QY 4 CPWDGHLHYREDQSPAPGRLCLNTMDAOSGLASAPVS-----GAGNHSYCRNPDE 55
DB 286 CPEGGITRYGTWSTASGAECLINW--NSSALSKPYARRPNKIKGLGNHYCRNPDR 344
QY 56 DPRGMCVYSGEAGVPEKRPCEDLRCPE-----TTSQALPAPFTTIEQDASBFG 101
DB 345 DPK-PWCYVP-KAKGKTYTFCSTPACPKPTEDCVYGVKGYTRGTHSFTT--SKASCLFW 237

Query Match 11.4%; Score 146.5; DB 1; Length 655;
Best Local Similarity 36.9%; Pred. No. 0,00014;
Matches 41; Conservative 10; Mismatches 41; Indels 19; Gaps 6;

QY 4 CPWDGHLHYREDQSPAPGRLCLNTMDAOSGLASAPVS-----GAGNHSYCRNPDE 55
DB 286 CPEGGITRYGTWSTASGAECLINW--NSSALSKPYARRPNKIKGLGNHYCRNPDR 344
QY 56 DPRGMCVYSGEAGVPEKRPCEDLRCPE-----TTSQALPAPFTTIEQDASBFG 101
DB 345 DPK-PWCYVP-KAKGKTYTFCSTPACPKPTEDCVYGVKGYTRGTHSFTT--SKASCLFW 237

A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-291 <SIE>
A:Cross-references: EMBL:X13097; NID:g35282; PIDD:CAA1489.1; PID:g35283
C:Comment: For the main splice form, see PIR:UKHUT. This form probably does not have pro
C:Genetics:
A:Gene: GDB:PLAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-8p12
A:Intron: 24/3: 39/1: 65/1: 122/1: 180/2: 211/1: 268/2
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: alternative splicing; fibrinolytic; glycoprotein; kringle
F1-23/Domain: signal sequence #status predicted <SIG>
F1-24-32/Domain: propeptide #status predicted <PRO>
F1-33-39/Product: t-plasminogen activator, inactive endothelial splice form #status pred
F1-41-78/Domain: fibronectin type I repeat homology <EFG>
F1-86-119/Domain: EGF homology <KRI>
F1-127-208/Domain: kringle homology #status atypical <KR2>
F1-215-291/Domain: kringle homology #status atypical <KR2>
F1-41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203/Dissulfide bonds: #status pred
Query Match 11.3%; Score 145.5; DB 2; Length 291;
Best Local Similarity 38.5%; Pred. No. 6;9e-05;
Matches 35; Conservative 6; Mismatches 37; Indels 13; Gaps 4;
Qy 4 CFWDNGHYRBDQSPAPGRLCNWLDAGSLASAPVS-----GAGNHSYCRNPE 54
Db 127 CYEDQGISYRGWSTAESGAECTW--NSRSLAQNAYSGRPDAIRLGLGNHNCYRPR 184
Qy 55 DRGPMCVYSGAGVPEKPECDRCPEPTTS 85
Db 185 DSK-FMCTVF-KAGKVSSEFCSPACSEGN 213
RESULT 6
PLU plasmin (EC 3.4.21.7) precursor [validated] - human
N:Alternate names: plasminogen precursor [mismom]
N:Contains: angiotatin; microplasmin; plasminogen
C:Species: Homo sapiens (man)
C:Date: 24-Apr-1984 #sequence revision 02-Dec-1994 #text change 15-Sep-2000
C:Accession: A35229; I52242; A26646; 163738; 184609; S03735; A00929; A04627; A04625; A04
J:Biochem. 265: 6104-6111, 1990
A:Title: Characterization of the gene for human plasminogen, a key proenzyme in the fibr
A:Reference number: A35229; PMID:90202879; PMID:2318848
A:Accession: A35229
A:Molecule type: DNA
A:Residues: 1-810 <PPT>
A:Cross-references: GB:U05286; GB:M34276; NID:g190064; PIDD:AAA60113.1; PID:g387026
A:Experimental source: leukocyte; lung fibroblast
R:Waldarrett, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Meroni, G.; Ottolenghi, S.; T
Biochem. Biophys. Res. Commun. 173: 1013-1018, 1990
A:Title: Definition of the transcription initiation site of human plasminogen gene in 1
A:Reference number: I52242; PMID:91097553; PMID:2266308
A:Accession: I52242
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-16 <MAL>
A:Cross-references: GB:M62890; NID:g190092; PIDD:AAA36454.1; PID:g553613
R:Forstgen, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.
FEBS Lett. 213: 254-260, 1987
A:Title: Molecular cloning and characterization of a full-length cDNA clone for human pl
A:Reference number: A26646; PMID:87162490; PMID:3030813
A:Accession: A26646
A:Molecule type: mRNA
A:Residues: 1-471; 'D', 473-810 <FOR>
A:Cross-references: GB:X05199; NID:g35530; PIDD:CAA28831.1; PID:g35531
A:Experimental source: liver
R:Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23: 4243-4250, 1984
A:Title: Characterization of a complementary deoxyribonucleic acid coding for human and
A:Reference number: I45961; PMID:85023311; PMID:6148961

A:Accession: I62738
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 292-471; 'D', 473-810 <MAL>
A:Cross-references: GB:K02922; NID:g190112; PIDD:AAA60124.1; PID:g387031
A:Accession: I64609
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 367-419 <ML3>
A:Cross-references: GB:K02921; NID:g190110; PIDD:AAA60123.1; PID:g190111
R:Brustholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergler, W.; Manneberg, M
Eur. J. Biochem. 114: 465-470, 1981
A:Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,
A:Reference number: S03735; PMID:8121097; PMID:7238497
A:Accession: S03735
A:Molecule type: Protein
A:Residues: 20-71; 'E', 73-76 <BRU>
R:Sortrup-Jensen, L.; Petersen, T.E.; Magnusson, S.
submitted to the Atlas, July 1977
A:Reference number: A00929
A:Accession: A00929
A:Molecule type: Protein
A:Residues: 20-71; 'E', 73-85, 87-106, 'D', 108-360, 'E', 362-810 <SOT>
R:Wiman, B.
Eur. J. Biochem. 76: 129-137, 1977
A:Title: Primary structure of the B-chain of human plasmin.
A:Reference number: A04627; PMID:7725245; PMID:142009
A:Accession: A04627
A:Molecule type: Protein
A:Residues: 581-810 <WII>
R:Wiman, B.; Wallen, P.
Eur. J. Biochem. 50: 469-494, 1975
A:Title: Structural relationship between "glutamic acid" and "lysine" forms of human pla
A:Reference number: A04625; PMID:75093329; PMID:122932
A:Accession: A04625
A:Molecule type: Protein
A:Residues: 20-50; 'Q', 51-71, 'E', 73-85, 87-100 <W12>
R:Wiman, B.; Wallen, P.
Eur. J. Biochem. 58: 539-547, 1975
A:Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen tha
A:Reference number: A04626; PMID:76043692; PMID:126863
A:Accession: A04626
A:Molecule type: Protein
A:Residues: 483-507; 'E', 509-604 <W13>
R:Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summatta, L.
J. Biol. Chem. 248: 1631-1633, 1973
A:Title: The primary structure of human plasminogen. II. The histidine loop of human pl-
A:Reference number: A92125; PMID:73149248; PMID:4694729
A:Contents: annotation; active site
R:Groekopf, W.R.; Summatta, L.; Robbins, K.C.
J. Biol. Chem. 244: 3590-3597, 1969
A:Title: Studies on the active center of human plasmin. Partial amino acid sequence of a
A:Reference number: A92048; PMID:69234739; PMID:4240117
A:Contents: annotation; active site
R:Textler, M.; Vail, Z.; Pecthy, L.
J. Biol. Chem. 257: 7401-7406, 1982
A:Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen.
A:Reference number: A92182; PMID:82213905; PMID:6919539
A:Contents: annotation; omega-aminocarboxylic acid binding sites
R:Vail, Z.; Pecthy, L.
J. Biol. Chem. 259: 13690-13694, 1984
A:Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential
A:Reference number: A92458; PMID:85054794; PMID:6094526
A:Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site
R:Caio, Y.; Uli, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehndel, S.; McCance, S.G.;
J. Biol. Chem. 271: 29461-29467, 1996
A:Title: Kringle domains of human angiotensin. Characterization of the anti-proliferativ
A:Reference number: A58813; PMID:97067211; PMID:8910613
A:Contents: annotation
R:Jensen, H.R.; Uguw, F.; Bini, A.; Colten, D.
Biochemistry 37: 4699-4702, 1998
A:Title: Generation of an angiotensin-like fragment from plasminogen by streptolysin-1 (M

A;Reference number: A58012; MUID:9548733; PMID:9548733
A;Contents: annotation
R;Tulinsky, A.; Mulichak, A.M.
submitted to the Brookhaven Protein Data Bank, July 1991
A;Reference number: A51341; PDB:1PK4
A;Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454
R;Tulinsky, A.; Wu, T.P.
submitted to the Brookhaven Protein Data Bank, July 1991
A;Reference number: A51488; PDB:2PK4
A;Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 375-454
R;Wu, T.P.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, August 1993
A;Reference number: A51911; PDB:1PKR
A;Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-181
R;Padmanabhan, K.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, April 1994
A;Reference number: A52408; PDB:1PKK
A;Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-454
R;Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A;Reference number: A65244; PDB:1CEA
A;Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R;Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A;Reference number: A65245; PDB:1CEB
A;Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R;Mulichak, A.M.; Tulinsky, A.; Ravichandran, K.G.
Biochemistry 30, 10576-10588, 1991
A;Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.9 Å
A;Reference number: A58819; MUID:92031502; PMID:1657148
A;Contents: annotation
R;Wu, T.P.; Padmanabhan, K.; Tulinsky, A.; Mulichak, A.M.
Biochemistry 30, 10589-10594, 1991
A;Title: The refined structure of the epsilon-aminocaproic acid complex of human plasmin
A;Reference number: A58818; MUID:92031503; PMID:1657149
A;Contents: annotation
A;Contents: annotation
R;De Voe, A.M.; Ultsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Westbrook, M.
Biochemistry 31, 270-279, 1992
A;Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator at 2.4 Å
A;Reference number: A39483; MUID:92118803; PMID:1310033
A;Contents: annotation; X-ray crystallography, 2.4 angstroms
R;Stein, B.; Teeter, M.M.; Whitlow, M.; Yamano, A.
submitted to the Brookhaven Protein Data Bank, June 1995
A;Reference number: A65980; PDB:1KRN
A;Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-454
R;Rejzante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A;Reference number: A65803; PDB:1HPJ
A;Contents: annotation; conformation by (1)H-NMR, residues 103-181
R;Rejzante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A;Reference number: A65804; PDB:1HPK
A;Contents: annotation; conformation by (1)H-NMR, residues 103-181
R;Rejzante, M.; Llinas, M.
Eur. J. Biochem. 221, 927-937, 1994
A;Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 1.
A;Reference number: S43645; MUID:94237157; PMID:8181475
A;Contents: annotation; conformation by (1)H-NMR, residues 96-184.
R;Rejzante, M.R.; Llinas, M.
Eur. J. Biochem. 221, 939-949, 1994
A;Title: Solution structure of the epsilon-aminocaproic acid complex of human plasminogen
A;Reference number: A58817; MUID:94237158; PMID:8181476
A;Contents: annotation; conformation by (1)H-NMR
C;Comment: Plasminogen is synthesized by the kidney and is present in plasma and many other
C;Comment: Plasminogen is converted to plasmin by plasminogen activators (see PIR:UKHU a
d PIR:FGHUGB).
C;Comment: Plasmin is inactivated by alpha-2-antiplasmin (see PIR:ITHU2) immediately after
rg-580, resulting in two chains connected by two disulfide bonds. Without the inhibitor
C;Comment: Microplasma is formed by autolytic cleavage of plasmin under artificial cond
ting solid tumors.
C;Genetic: GDB:PLG
C;Genetics:

A:Cross-references: GDB:119498; OMIM:173350
A:Map position: 6q26-q27
A:Introns: 17/1; 62/2; 98/1; 136/2; 183/1; 223/2; 263/1; 317/2; 366/1; 419/2; 480/1; 525

C:Function:
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a variety of tissues; also activates the urokinase-type plasminogen activator on the walls of the glomerular follicle;
A:Pathway: fibrinolysis
C:Superfamily: plasmin; kringe homology; plasminogen-related protein precursor homology
C:Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; hydrolysis
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-810/Product: plasminogen #status experimental <PRO>
F:20-96/Domain: activation peptide #status experimental <APR>
F:79-466/Product: angiotensin #status experimental <AST>
F:97-580,581-810/Product: plasmin #status experimental <MAT>
F:103-181/Domain: plasmin chain A #status experimental <CHA>
F:185-262/Domain: kringe homology <KR1>
F:275-352/Domain: kringe homology <KR2>
F:377-454/Domain: kringe homology <KR4>
F:481-560/Domain: kringe homology <KR5>
F:550-580,581-810/Product: microplasmin #status experimental <MMT>

Query Match 11.0%; Score 142; DB 1; Length 810;
Best Local Similarity 37.2%; Pred. No. 0.00044;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

Oy 4 CFWDNGHLYREDQTSPPAGRLCLNWLDA-----OSGLASAPVSGAGNHSCRNPDDEDPRG 58
Db 103 CKTGNGKXRYGRGMSKTNGKIGICQKMSSSRPHRPSPATHPSEEL--ENYCRNPDDNDPQG 161
Oy 59 PWCYVSGEAGVPEKR--PCEDLRCPD 82
Db 162 PWCYTDD---PEKRYDCDIIECEE 183

RESULT 7
A:Accession: A40522
plasma (EC 3.4.21.7) precursor - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 28-Feb-1992 #sequence_revision 17-Apr-1993 #text_change 16-Jul-1999
C:Accession: A40522
R.Kanala, U.J.; Makker, S.P.
J. Biol. Chem. 266, 10825-10829, 1991
A>Title: Identification of the rat Heymann nephritis autoantigen (GP330) as a receptor for anti-
A:Reference number: A40522; PMID:91250378; PMID:1645711
A:Accession: A40522
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-169 <KAN>
A:Cross-references: GB:M62832; NID:g206215; PIDN:AAA41884.1; PID:g554488
A>Note: The authors translated the codon TCT for residue 76 as Ala
C:Superfamily: plasmin; kringe homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringe; serine proteinase
F:j34-112/Domain: kringe homology <KR>
F:j34-112,55-95,83-107/Disulfide bonds: #status predicted

Query Match 10.7%; Score 138; DB 2; Length 169;
Best Local Similarity 31.8%; Pred. No. 0.00016;
Matches 35; Conservative 13; Mismatches 48; Indels 14; Gaps 5;

Oy 4 CFWDNGHLYREDQTSPPAGRLCLNWLDA-----OSGLASAPVSGAGNHSCRNPDDEDPRG 58
Db 34 CYGNGKSXYGRHSSTNNKKCCQSWWSMTPIHSHSKTPANFPDSGL--ENMYCRNPDDNDORG 92
Oy 59 PWCYVSGEAGVPEKR--PCEDLRCPETTSQALPARPTTIQASBEGADE 106
Db 93 PWCFTTD---PSVWEYCNIKRCEGTGGV--AESAIPOVPAPSAYISE 136

RESULT 8
u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon

C.Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)
C.Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999
C.Accession: S14687, S08651
R.Au, Y.F.T.; Wang, T.W.; Clowes, A.W.
Nucleic Acids Res. 18, 3411, 1990
A.Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen
A.Reference number: S14687; MUID:90287734; PMID:2113276
A.Accession: S14687
A.Molecule type: mRNA
A.Residues: 1-433 <AU>
A.Cross-references: EMBL:X51935; NID:G38130; PIDN:CAA36200.1; PID:G38131
C.Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C.Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F.1-40/Domain: signal sequence #status predicted <STG>
F.21-176/Product: plasminogen activator chain A #status predicted <ACH>
F.30-61/Domain: EGF homology <EGF>
F.69-150/Domain: Kringle homology <KRG>
F.178-433/Product: plasminogen activator chain B #status predicted <BCH>
F.178-421/Domain: trypsin homology <TRY>
F.167-298, 208-324, 216-287, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted
F.323, 274, 378/Active site: His, Asp, Ser #status predicted
F.332/Binding site: carbonylate (Asn) (covalent) #status predicted

| | | | | |
|--------------------------|--------|--------------------|------------|-------------|
| Query Match | 10.7%; | Score 138; | DB 1; | Length 433; |
| Best Local Similarity | 32.7%; | Pred. No. 0.00047; | | |
| Matches 32; Conservative | 14; | Mismatches 36; | Indels 16; | Gaps 4; |

QY 4 CFMDNGHLYREQTSPAGRLCLNLDA-----QSGLASPVSGAGNHSYCNPDEDP 56
| : ||||| :: | | | | : | | | | :
Db 69 CYEGNGHFYRGKASTDITWGRSCLAWNSATVLQQTTHAHRSDALQLGELKHNYCNPDI-NR 122

```

QY 57 RGPWCVSGEAGVPEK-----RPECDLRCPETTSAL 88
      ||||| : : : : :
Db 128 RRPWCYV--QYGLKQRYGECMHNACADKKPSSPEEL 163

```

RESULT 9

U-plasminogen activator (EC 3.4.21.73) precursor - pig
N/Alternate names: uPA
C/Species: Sus scrofa domestica (domestic pig)
C/Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
C/Accession: A00932
R/Ngamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
N/Nucleic Acids Res. 12, 9525-9541, 1984
A/Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
A/Reference number: A00932; MUID:85087954; PMID:6096832
A/Accession: A00932
A/Molecule type: DNA
A/Residues: 1-240, 'H', 242-442 <NAG1>
A/Experimental source: kidney cell line LLC-PK1
R/Ngamine, Y.
submitted to the Protein Sequence Database, December 1986
A/Reference number: A37566
A/Contents: annotation; correction to residue 241
C/Genetics:
A/Intons: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
C/Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C/Keywords: glycoprotein, heterodimer; hydrolase; kringle; serine proteinase
F/1-20/Domain: signal sequence #status predicted <SIG>
F/21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F/33-64/Domain: EGF homology <EGF>
F/72-153/Domain: kringle homology <KR>
F/190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F/190-430/Domain: trypsin homology <TRY>
F/152/Binding site: carbohydrate (Asn) (covalent) #status predicted
F/119-310-220-236-228-229,324-393,356-372,383-411/Disulfide bonds: #status predicted
F/25,286,387/Active site: His, Asp, Ser #status predicted

| | | | | |
|-----------------------|--------------|--------------------|----------------|-------------|
| Query Match | 10.5% | Score 135.5; | DB 1; | Length 442; |
| Best Local Similarity | 36.9%; | Pred. No. 0.00079; | | |
| Matches 38; | Conservative | 7; | Mismatches 37; | Indels 21; |
| | | | | Gaps 5 |

QY 4 CFMDNGHLYRSDQTSPPAPGLRCLNWLDAOSGL-----ASAPVS---GAGNSYCNRPEDP 56
Db 72 CFEQNGHSHYRKAQNTNNGRGRCLPMSATVLTNLYHHRRDPAQLGLGKKNYCNRPD-NQ 130
QY 57 RGPWCYVS-----GEAGVP-----EKRPCELRCPETTSQ 86
Db 131 RRPWCYVAVGLKQLVDEQWPNCSGGSHRPAYVGNKRPFTPE 173

RESULT 10

E-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat
 N:Alternate names: tissue plasminogen activator
 C:Species: *Desmodus rotundus* (common vampire bat)
 C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C/Accession: U50599
 R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boldol, W.; Klingmann, P.; Alagon, A.; Don
 Gene 105, 229-237, 1991
 A>Title: The plasminogen activator family from the salivary gland of the vampire bat *Des*
 A:Reference number: U50597; PMID:92039036; PMID:1937019
 A:Accession: J50599
 A:Molecule type: mRNA
 A:Residue: 1-431 <KKA>
 A:Cross-references: GB:J63389; NID:g166076; PID:AAA31594.1; PID:g166077
 A:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F:1.1/Domain: signal sequence #status predicted <SIG>
 F:12-36/Domain: propeptide #status predicted <PRO>
 F:37-43/Domain: EGF homology <EGF>
 F:174-74/Domain: EGF homology <EGF>
 F:82-163/Domain: kringle homology <KRG>
 F:180-425/Domain: tryptsin homology <TRY>
 F:41-52,46-63,65-74,82-163,103-145,134-158,168-299,211-227,219-288,313-368/Disulfide bon
 F:119-352/Binding site: carbonylhydrate (Asn) (covalent) #status predicted
 F:179-180/Cleavage site: His-Ser (plasmin) #status predicted
 F:226,275,382/Active site: His, Asp, Ser #status predicted
 F:345-361,378-406/Disulfide bonds: #status predicted

| | | | | |
|--------------------------|-------|-----------------|-----------|------------|
| Query Match | 10.4% | Score 134 | DB 2 | Length 431 |
| Best Local Similarity | 38.6% | Pred. No. 0.001 | | |
| Matches 27; Conservative | 9 | Mismatches 20 | Indels 14 | Gaps 3 |

| | | | | |
|-----|---|----|---|-----|
| OY | | 4 | C F W D N G H L Y R E D Q T S A P G I R C L N W I D A O S G I-----A S A P V S G A C H S Y C N P D | 53 |
| | : | : | : : : | : |
| D b | | 82 | C Y K D Q G T Y T R G M S T E S G A C I N W---N S U L T R K T Y N G R S D A L T G L G N H Y C N P D | 138 |
| | : | : | : : : | : |

| | | | |
|----|-----|------------|-----|
| Qy | 54 | EDPRGPMCTV | 63 |
| | : | | |
| Db | 139 | NNSK-PMCTV | 147 |

RESULT 11

t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
C/Species: Megaderma lyra
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C/Accession: A34369
J. Biold. S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jacob
R. Biold. Chem. 264, 17947-17952, 1969
A./Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasmin
A./Reference number: A34369; PMID:90036667; PMID:2509450
A./Accession: A34369
A./Status: preliminary
A./Molecule type: mRNA
A./Residues: 1-477 <GAR>
A./Cross-references: NID:G166080; PIDN:AA31596.1; PID:G166081
C/Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C/Keywords: fibrinolysis; glycoprotein; hydroxylase, kringle, serine proteinase
F./1-21/Domains: signal sequence #status predicted <PRO> #SIGs
F./22-36/Domains: propeptide #status predicted <PRO> #SIGs
F./37-477/Product: plasminogen activator #status predicted <PLA>
F./42-79/Domains: fibronectin type I repeat homology <1FA>
F./87-120/Domains: EGF homology <EGF>

A:Residues: 66-431 <VER>
A:Cross-references: GB:D00244; NID:g220138
R:Jacob, P.; Cravador, A.; Lortiau, R.; Brockly, F.; Colau, B.; Chuchana, P.; Van Elsen, DNA 4, 139-146, 1985
A:Title: Molecular cloning, sequencing, and expression in *Escherichia coli* of human prepro-urokinase
A:Reference number: 138102; MUID:8520359; PMID:3888571
A:Accession: 138102
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-150, 'W', 152-213, 'T', 215-385, 'C', 387-429, 'V', 431 <TAC>
A:Cross-references: EMBL:X02760; NID:935297; PIDN:DA46535.1; PID:G35298
R:Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, Biochim. Biophys. Acta 1293, 83-89, 1996
A:Title: Characterization of single chain urokinase-type plasminogen activator with a novel propeptide
A:Reference number: 865783; MUID:96186279; PMID:8652631
A:Accession: 865783
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 21-140, 'L', 142-213, 'T', 215-431 <YOS>
A:Cross-references: EMBL:D1143; NID:91311467; PIDN:BA01919.1; PID:g1199928
R:Günzler, W.A.; Steffens, G.J.; Oetting, F.; Kim, S.M.A.; Frankus, E.; Flohe, L. Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982
A:Title: The primary structure of high molecular mass urokinase from human urine.
A:Reference number: A37562; MUID:83055084; PMID:6754569
A:Accession: A37562
A:Molecule type: protein
A:Residues: 21-177 <GN>
R:Schaller, J.; Nick, H.; Rickli, E.E.; Gillesen, D.; Lergier, W.; Studer, R.O. Eur. J. Biochem. 125, 251-257, 1982
A:Title: Human low-molecular-weight urinary urokinase. Partial characterization and preliminary amino acid sequence
A:Reference number: A37563; MUID:83003608; PMID:6749491
A:Accession: A37563
A:Molecule type: protein
A:Residues: 156-176, 179-193, 'T', 195, 'T', 197-224 <SCH>
R:Steffens, G.J.; Günzler, W.A.; Oetting, F.; Frankus, E.; Flohe, L. Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982
A:Title: The complete amino acid sequence of low molecular mass urokinase from human urine
A:Reference number: A37564; MUID:83055099; PMID:6754572
A:Accession: A37564
A:Molecule type: protein
A:Residues: 158-410 <SPB>
R:Kentzer, E.J.; Buko, A.; Menon, G.; Sarin, V.K. Biochem. Biophys. Res. Commun. 171, 401-406, 1990
A:Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinant urokinase
A:Reference number: A35689; MUID:90365737; PMID:2393398
A:Accession: A35689
A:Molecule type: protein
A:Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <EN>
A:Note: Identification of a fucose and attempt to determine its attachment site
R:Rabdan, S.A.; Desjardins, J.; Bell, A.W.; Barrville, D.; Mazar, A.; Henkin, J.; Goltz, Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990
A:Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell line
A:Reference number: A36697; MUID:91097529; PMID:2125213
A:Accession: A36697
A:Molecule type: protein
A:Residues: 21-34 <RAB>
R:Li, X.; Bokman, A.M.; Linas, M.; Smith, R.A.G.; Dobson, C.M. submitted to the Brookhaven Protein Data Bank, July 1993
A:Reference number: A51255; PDB:1XDU
A:Contents: annotation; conformation; and disulfide bond assignments by (1)H-NMR, residue R:Li, X.; Smith, R.A.G.; Dobson, C.M. Biochemistry 31, 9582-9577, 1992
A:Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain of human urokinase
A:Reference number: A44375; MUID:93003110; PMID:1327118
A:Contents: annotation; conformation; and disulfide bond assignments by (1)H-NMR R:Hanen, A.P.; Petros, A.M.; Meadows, R.P.; Nettesheim, D.G.; Mazar, A.P.; Olejniczak, submitted to the Brookhaven Protein Data Bank, January 1994
A:Reference number: A66822; PDB:1URK
A:Contents: annotation; conformation; and disulfide bond assignments by (1)H-NMR, residue R:Sprague, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.; submitted to the Brookhaven Protein Data Bank, July 1995
A:Reference number: A66058; PDB:1LMW
A:Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175, 179-426

C/Comment: This enzyme is found in urine in a high molecular mass form, consisting of A

C/Comment: urokinase-type plasminogen activator proteolytically activates plasminogen, a

C/Genetics:

A:Gene: GDB:PMU

A:Cross-references: GDB:119497; OMIM:191840

A:Map position: 10q24-10q24

A:Positions: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3

C/Function:

A>Description: proteolytically activates plasminogen

A:Pathway: fibrinolysis

C/Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C/Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine proteina

F:1-20/Domin: signal sequence #status predicted <Sig>

F:21-43/Product: urokinase-type plasminogen activator, single chain form #status predic

F:21-17/Product: urokinase-type plasminogen activator chain A #status experimental <MPA

F:31-62/Domin: EGF homology <EGF>

F:70-151/Domin: kringle homology <KR>

F:178-177/Product: urokinase-type plasminogen activator chain B #status experimental <MP

F:179-419/Domin: trypsin homology <TRY>

F:31-39; 33-51; 53-62; 70-151; 91-133; 122-146; 168-299; 209-225; 217-288; 313-382; 345-361; 372-40

F:38/Binding site: carbohydrate (Thr) (covalent) #status predicted

F:178-179/Cleavage site: Lys-116 (plasmin) #status experimental

F:224; 275; 376/Active site: His, Asp, Ser #status experimental

F:322/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 10.28; Score 132; DB 1; Length 431;

Best Local Similarity 32.78; Pred. No. 0.0015;

Matches 33; Conservative 13; Mismatches 33; Indels 22; Gaps 5;

QY 57 RGPWCYSGEAGVDEKRP-----CEDLRCEPTTSQAL 88

Db 70 CYEGNGHYRSDQSPAPGLRCLNWLDA-----QSGLASAPVSGAGNSYCRNPDEDP 56

Db 156 CYEGNGHYRSGKASTDTPGRCLPMSNATVYQOTYHHRSDALDGLGKINYNCRNP-NR 128

Db 129 RRPWCYV--QVGL--KPLVGECHWDCADEKSKSPPEEL 164

RESULT 15

A40332

macrophage-stimulating protein 1 precursor - mouse

N.Alternate names: hepatocyte growth factor-like protein

C.Species: Mus musculus (house mouse)

C.Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 18-Jun-1999

C.Accession: A40332; B40332

R.Degen, S.J.F.; Stuart, L.A.; Han, S.; Jamison, C.S.

Biochemistry 30, 9781-9791, 1991

A.Title: Characterization of the mouse cDNA and gene coding for a hepatocyte growth factor

A.Reference number: A40332; MUID:92002017; PMID:1832957

A.Accession: A40332

A.Molecule type: DNA

A.Residues: 1-716 <DEG>

A:Cross-references: GB:M74180; NID:g193831; PIDN:AAA50166.1; PID:g193832

A:Accession: B40332

A.Molecule type: mRNA

A.Residues: 1-18; P'20-716 <DEG2>

A:Cross-references: GB:M74181; NID:g193833; PIDN:AAA50167.1; PID:g193834

C.Genetics: 18/1; 67/2; 105/1; 143/2; 189/1; 229/2; 269/1; 334/2; 378/1; 412/2; 458/1; 47

C/Comment: disulfide-bonded heterodimer of chains derived from the same precursor

C/Superfamily: hepatocyte growth factor; kringle homology; trypsin homology

C/Keywords: duplication; glycoprotein; growth factor; kringle

F:1-31/Domin: signal sequence #status predicted <Sig>

F:19-488; 489-716/Product: macrophage-stimulating protein 1 #status experimental <MAT>

F:19-483/Domin: alpha chain #status experimental <ACH>

F:110-186/Domin: kringle homology <KR1>

F:191-268/Domin: kringle homology <KR2>

F:292-370/Domin: kringle homology <KR3>

F:379-457/Domin: kringle homology <KR4>

F:484-711/Domin: beta chain #status experimental <BCH>

F:489-709/Domin: trypsin homology <TRY>

F:72; 173; 305; 620/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

Query Match      10.1%; Score 130.5; DB 1; Length 716;
Best local similarity 30.7%; Pred. No. 0.0037;
Matches 35; Conservative 7; Mismatches 37; Indels 35; Gaps 5;

Qy 4 CFWDNGHLVREDQTSPPGLRCLNW---LDAQSLASAPVSGAGNHSYCRNPDEDPRGFW 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 110 CIMDNGVSYRGTVARTAGILPCQAMSRPRPNDHKYTPPKNGL-EENFCRNPDGDPGRGFW 168
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 61 CYVS-----GEAGVPEK-RPCE--DLRCPET 83
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 169 CYTTRRSVRFQSGIKTCREAVCVLQNGEDYRGEDVVTESGRCQRWDLQHPHS 222
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: April 7, 2003, 09:19:19
 Job time : 21.6216 secs

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:24 ; Search time 11.0686 Seconds

(without alignments)
906.823 Million cell updates/sec

Title: US-10-057-951-2_COPY_22_263

Sequence: 1 SGSGCFWMDNGLHYREDQTPA.....PVDQEGSTPLMGAGTRQA 242

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 156.5 | 12.1 | 566 | 1 | TPA BOVIN |
| 2 | 154.5 | 12.0 | 562 | 1 | TPA_HUMAN |
| 3 | 154 | 11.9 | 559 | 1 | TPA_RAT |
| 4 | 150 | 11.6 | 653 | 1 | HGF_MOUSE |
| 5 | 147.5 | 11.4 | 559 | 1 | HGF_MOUSE |
| 6 | 146.5 | 11.4 | 655 | 1 | HGF_MOUSE |
| 7 | 142 | 11.0 | 810 | 1 | PLMN_HUMAN |
| 8 | 138 | 10.7 | 169 | 1 | PLMN_RAT |
| 9 | 138 | 10.7 | 433 | 1 | UROK_PAPCY |
| 10 | 135.5 | 10.5 | 442 | 1 | UROK_PIG |
| 11 | 134 | 10.4 | 431 | 1 | UROK_PIG |
| 12 | 134 | 10.4 | 477 | 1 | UROK_PIG |
| 13 | 132 | 10.2 | 431 | 1 | UROK_PIG |
| 14 | 130.5 | 10.1 | 716 | 1 | HGF_MOUSE |
| 15 | 129 | 10.0 | 810 | 1 | PLMN_HUMAN |
| 16 | 127 | 9.9 | 394 | 1 | UROK_PIG |
| 17 | 127 | 9.9 | 433 | 1 | UROK_PIG |
| 18 | 126.5 | 9.8 | 728 | 1 | HGF_MOUSE |
| 19 | 124.5 | 9.7 | 810 | 1 | PLMN_HUMAN |
| 20 | 124 | 9.6 | 622 | 1 | THRB_HUMAN |
| 21 | 124 | 9.6 | 728 | 1 | HGF_MOUSE |
| 22 | 124 | 9.6 | 812 | 1 | PLMN_HUMAN |
| 23 | 123.5 | 9.5 | 434 | 1 | UROK_PIG |
| 24 | 123 | 9.5 | 593 | 1 | UROK_PIG |
| 25 | 122.5 | 9.5 | 4548 | 1 | UROK_PIG |
| 26 | 122 | 9.5 | 603 | 1 | UROK_PIG |
| 27 | 121.5 | 9.4 | 728 | 1 | HGF_MOUSE |
| 28 | 120.5 | 9.3 | 333 | 1 | PLMN_HUMAN |
| 29 | 120 | 9.3 | 790 | 1 | PLMN_HUMAN |
| 30 | 119.5 | 9.3 | 477 | 1 | UROK_PIG |
| 31 | 119 | 9.2 | 433 | 1 | UROK_PIG |
| 32 | 118.5 | 9.2 | 615 | 1 | UROK_PIG |
| 33 | 118 | 9.2 | 432 | 1 | UROK_PIG |

| | | | | | | |
|----|-------|-----|------|---|------------|-----------------------|
| 34 | 117 | 9.1 | 625 | 1 | THRB_BOVIN | P00735 bos taurus |
| 35 | 116 | 9.0 | 1420 | 1 | APOL_MACMU | P14417 macaca mulatta |
| 36 | 115.5 | 9.0 | 711 | 1 | HGF_HUMAN | P26927 homo sapien |
| 37 | 114.5 | 8.9 | 812 | 1 | PLMN_MOUSE | P20918 mus musculus |
| 38 | 114 | 8.8 | 473 | 1 | KREM_MOUSE | Q99n43 mus musculus |
| 39 | 114 | 8.8 | 473 | 1 | KREM_MOUSE | Q92464 ratius norv |
| 40 | 114 | 8.8 | 475 | 1 | KREM_MOUSE | Q96m08 homo sapien |
| 41 | 114 | 8.8 | 475 | 1 | KREM_MOUSE | P19221 mus musculus |
| 42 | 109.5 | 8.5 | 1709 | 1 | SN_HUMAN | Q9b222 homo sapien |
| 43 | 107 | 8.3 | 617 | 1 | THRB_RAT | P18292 ratius norv |
| 44 | 103.5 | 8.0 | 343 | 1 | PLMN_SHEEP | P81286 ovis aries |
| 45 | 101.5 | 7.9 | 325 | 1 | PLMN_PETMA | P33574 petromyzon |

ALIGNMENTS

```

RESULT 1
ID TPA_BOVIN STANDARD; PRT; 566 AA.
AC Q26198;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
and tPA."
RL Int. Dairy J. 5:605-617(1995).
CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
MANY OTHER PHYSIOLOGICAL EVENTS.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
plasminogen to form plasmin.
CC -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
BOND.
CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X85800; CAA59795.1; -.
DR HSSP; P00750; IRTF.
DR MEROPS; S01.232; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000083; Fibrinectn1.

```

```
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; TRYPSIN; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Plasma; Kringler; EGF-like domain; Repeat; Signal.
KM Plasma; Kringler; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 21
FT PROPEP 22 33
FT CHAIN 34 566
FT CHAIN 34 314
FT CHAIN 315 566
FT CHAIN 315 566
FT DOMAIN 40 82
FT DOMAIN 83 121
FT DOMAIN 128 209
FT DOMAIN 219 300
FT DOMAIN 315 566
FT ACT_SITE 361 361
FT ACT_SITE 410 410
FT ACT_SITE 517 517
FT ACT_SITE 517 517
FT DISULFID 42 72
FT DISULFID 70 79
FT DISULFID 87 98
FT DISULFID 92 109
FT DISULFID 111 120
FT DISULFID 128 209
FT DISULFID 149 191
FT DISULFID 180 204
FT DISULFID 219 300
FT DISULFID 240 282
FT DISULFID 271 295
FT DISULFID 303 434
FT DISULFID 346 362
FT DISULFID 354 423
FT DISULFID 448 523
FT DISULFID 480 496
FT DISULFID 513 541
FT CARBOHYD 153 153
FT CARBOHYD 487 487
SQ SEQUENCE 566 AA; 63701 MW; 2EB6BB4B32276C3 CRC64;

Query Match 12.1%; Score 156.5; DB 1; Length 566;
Best Local Similarity 36.7%; Pred. No. 6.4e-06;
Matches 36; Conservative 10; Mismatches 39; Indels 13; Gaps 4;

OY 4 CFWDNGLHYREDQTSAPGLRCLNWLDAQSGLASAPV-----GAGNHSYCRNPDE 54
DB 128 CYDDQVAVRGTSTAESGACANM--NSSGLAMKPYSGRRPVAIRLGLGNHNYCRNPQ 185
OY 55 DPGKPCIVSGEAGVEKRCEDLRCPETTSQLLPAFT 92
DB 186 DSK-PWCYVF-KAGXYISEFCSTPACAKVAEBDDCT 221

RESULT 2 *
```

```
TPA_HUMAN
ID TPA_HUMAN STANDARD; PRT; 562 AA.
AC P00750; Q15103;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator) (Alteplase) (Retepase).
GN PLAT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Melanoma;
RA MEDLINE=83115262; PubMed=6337343;
RA Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vohar G.A.,
RA Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,
RA Goeddel D.V., Collen D.;
RT "Cloning and expression of human tissue-type plasminogen activator
RT cDNA in E. coli.";
RL Nature 301:214-221 (1983).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Fetal lung;
RA MEDLINE=88262579; PubMed=3133640;
RA Sasaki H., Saito Y., Hayaishi M., Otsuka K., Niwa M.;
RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA
RT from human fetal lung cells.";
RL Nucleic Acids Res. 16:5695-5695 (1988).
RN [3]
RP SEQUENCE FROM N.A.
RA MEDLINE=88054470; PubMed=2824147;
RA Reddy V.B., Garramone A.J., Sasek H., Wei C.-M., Watkins P., Galli J.,
RA Hsing N.;
RT "Expression of human uterine tissue-type plasminogen activator in
RT mouse cells using BPV vectors.";
RL DNA 6:461-472 (1987).
RN [4]
RP SEQUENCE FROM N.A.
RA MEDLINE=86196143; PubMed=3009482;
RA Flierzer Degen S.O., Rajput B., Reich E.;
RT "The human tissue plasminogen activator gene.";
RL J. Biol. Chem. 261:6972-6985 (1986).
RN [5]
RP SEQUENCE FROM N.A.
RA MEDLINE=84298137; PubMed=6089198;
RA Ny T., Elgh F., Lund B.;
RT "The structure of the human tissue-type plasminogen activator gene:
RT correlation of intron and exon structures to functional and
RT structural domains.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359 (1984).
RN [6]
RP SEQUENCE FROM N.A.
RA MEDLINE=86284200; PubMed=3090401;
RA Harris T.J., Patel T., Marston F.A., Little S., Emage J.S.,
RA Odenaker G., Volckert G., Rombaux W., Billaud A., Somer P.;
RT "Cloning of cDNA coding for human tissue-type plasminogen activator
RT and its expression in Escherichia coli.";
RL Mol. Biol. Med. 3:279-292 (1986).
RN [7]
RP SEQUENCE OF 212-361 FROM N.A.
RA MEDLINE=83169656; PubMed=6572897;
RA Edlund T., Ny T., Raahy M., Heden L.-O., Palm G., Holmgren E.,
RA Josephson S.;
RT "Isolation of cDNA sequences coding for a part of human tissue
RT plasminogen activator.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352 (1983).
RN [8]
RP SEQUENCE OF 1-36 FROM N.A.
RA MEDLINE=85289338; PubMed=3161893;
RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,
```

RA Schlemming W.-D.; Isolation and characterization of the human tissue-type plasminogen
RT activator structural gene including its 5' flanking region.";
RL J. Biol. Chem. 260:11223-11230(1985).
RN [9]
RP SEQUENCE OF 31-562 FROM N.A.
RX MEDLINE=91291340; PubMed=1366681;
RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;
RT "Purification and characterization of tissue plasminogen activator
RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells.";
RL Agric. Biol. Chem. 55:1225-1232(1991).
RN [10]
RP SEQUENCE OF 36-562.
RC TISSUE=Melanoma;
RX MEDLINE=85000468; PubMed=6433976;
RA Poll G., Kaelin-Letroum M., Bergsdorf N., Wallen P., Joernvall H.;
RT "Tissue plasminogen activator: peptide analyses confirm an indirectly
RT derived amino acid sequence, identify the active site serine residue,
RT establish glycosylation sites, and localize variant differences.";
RL Biochemistry 23:3701-3707(1984).
RN [11]
RP SEQUENCE OF 33-52 AND 311-330.
RC TISSUE=Melanoma;
RX MEDLINE=83209620; PubMed=6682760;
RA Wallen P., Poll G., Bergsdorf N., Raaby M., Ny T., Joernvall H.;
RT "Purification and characterization of a melanoma cell plasminogen
RT activator.";
RL Eur. J. Biochem. 133:681-686(1983).
RN [12]
RP SEQUENCE FROM N.A. (SHORT ISOCORN).
RC TISSUE=Umbilical vein;
RX MEDLINE=90192129; PubMed=2107528;
RA Siebert P.D., Fong K.;
RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
RT human endothelial cells.";
RL Nucleic Acids Res. 18:1086-1086(1990).
RN [13]
RP STRUCTURE OF CARBOHYDRATES.
RX MEDLINE=90092112; PubMed=2513186;
RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;
RT "Carbohydrate structure of recombinant human uterine tissue
RT plasminogen activator expressed in mouse epithelial cells.";
RL Eur. J. Biochem. 186:273-286(1989).
RN [14]
RP CARBOHYDRATE-LINKAGE SITE THR-96.
RX MEDLINE=91159408; PubMed=1900431;
RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;
RT "Tissue plasminogen activator has an O-linked fucose attached to
RT chondron-61 in the epidermal growth factor domain.";
RL Biochemistry 30:2311-2314(1991).
RN [15]
RP DISULFIDE BONDS IN KRINGLE 2.
RX MEDLINE=9124765; PubMed=1645336;
RA Vlahos C.J., Wilhelm O.G., Hassell T., Jaskunas S.R., Bang N.U.;
RT "Disulfide pairing of the recombinant kringle-2 domain of tissue
RT plasminogen activator produced in Escherichia coli.";
RL J. Biol. Chem. 266:10070-10072(1991).
RN [16]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
RX MEDLINE=96200985; PubMed=8613982;
RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
RA Bode W.;
RT "The 2.3 A crystal structure of the catalytic domain of recombinant
RT two-chain human tissue-type plasminogen activator.";
RL J. Mol. Biol. 258:117-135(1996).
RN [17]
RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
RX MEDLINE=97449126; PubMed=9305622;
RA Renatus M., Eng R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,
RA Bode W.;
RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray
RT crystal structure of single-chain human tPA.";
RL EMBO J. 16:4797-4805(1997).

```

[18] X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.
RX MEDLINE=92118803; Pubmed=131003; Padmanabhan K., Tullinsky A.,
RA de Vos A., Ultsch M.H., Kelley R.F.,
RA Westbrook M.L., Kosciakof A.A.;
RT "Crystal structure of the kringle 2 domain of tissue plasminogen
RL activator at 2.4-A resolution."
RN Biochemistry 31:270-279(1992).
[19] STRUCTURE BY NMR OF KRINGLE 2.
RX MEDLINE=90122799; Pubmed=2558718;
RA Byeon I.-J.U., Kelley R.F., Llinas M.;
RT "1H NMR structural characterization of a recombinant kringle 2 domain
RL from human tissue-type plasminogen activator."
RN Biochemistry 28:9350-9360(1989).
[20] STRUCTURE BY NMR OF KRINGLE 2.
RX MEDLINE=91200042; Pubmed=1901789;
RA Byeon I.-J.U., Kelley R.F., Llinas M.;
RT "Kringle-2 domain of the tissue-type plasminogen activator. 1H-NMR
RL assignments and secondary structure."
RN Eur. J. Biochem. 197:155-165(1991).
[21] STRUCTURE BY NMR OF KRINGLE 2.
RX MEDLINE=92106329; Pubmed=1762144;
RA Byeon I.-J.U., Llinas M.;
RT "Solution structure of the tissue-type plasminogen activator kringle
RL 2 domain complexed to 6-aminohexanoic acid an antifibrinolytic
drug."
RN J. Mol. Biol. 222:1035-1051(1991).
[22] STRUCTURE BY NMR OF 38-85.
RX MEDLINE=92292163; Pubmed=1602484;
RA Downing A.K., Driscoll P.C., Harvey T.S., Dudgeon T.J., Smith B.O.,
RA Barton M., Campbell I.D.;
RT "Solution structure of the fibrin binding finger domain of
RL tissue-type plasminogen activator determined by 1H nuclear magnetic
RN resonance."
RJ J. Mol. Biol. 225:821-833(1992).
[23] STRUCTURE BY NMR OF 36-126.
RX MEDLINE=96027104; Pubmed=758289;
RA Smith B.O., Downing A.K., Driscoll P.C., Dudgeon T.J., Campbell I.D.;
RT "The solution structure and backbone dynamics of the fibronectin type
RL I and epidermal growth factor-like pair of modules of tissue-type
RN plasminogen activator."
RJ Structure 3:823-833(1995).
CC -I- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOPATHOLOGICAL EVENTS.
CC -I- CATALYTIC ACTIVITY: Specific cleavage of Arg-I-Val bond in
CC plasminogen to form plasmin.
CC -I- SUBUNIT: BINDS TO FIBRIN WITH HIGH AFFINITY. THIS INTERACTION
CC
QY Query Match 12.0%; Score 154.5; DB 1; Length 562;
DB Beat Local Similarity 39.6%; Pred.No. 9.4e-06;
QY Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
DB 55 DRGPGWCYSGRAGVPEKRPCEDLACPEPTTS 85
DB 127 CYEDGISIVRGWWSVAESGAECTNM--NSSALAOKPXSQRPDARLGLGNHYNCRNPDR 184
DB 185 DSK-PWCYVF-KAGKYSEFCSTPA CSGNS 213

```

DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (BC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89170114; PubMed=3148445;
RA NY T., Leonardsson G., Haueh A.J.W.;
RT "Cloning and characterization of a cDNA for rat tissue-type
RT plasminogen activator.";
RL DNA 7:671-677(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90130448; PubMed=2105315;
RA Feng P., Ohlsson M., Ny T.;
RT "The structure of the TATA-less rat tissue-type plasminogen activator
RT gene. Species-specific sequence divergences in the promoter predict
RT differences in regulation of gene expression.";
RL J. Biol. Chem. 265:2022-2027(1990).
CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED, EXTRACELLULAR.
CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALIKREIN OR FACTOR XA.
CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M23697, AAA41812.1, -;
DR EMBL, M31197, AAA42261.1, -;
DR EMBL, M31185, AAA42261.1, JOINED.
DR EMBL, M31186, AAA42261.1, JOINED.
DR EMBL, M31187, AAA42261.1, JOINED.
DR EMBL, M31188, AAA42261.1, JOINED.
DR EMBL, M31189, AAA42261.1, JOINED.
DR EMBL, M31190, AAA42261.1, JOINED.
DR EMBL, M31191, AAA42261.1, JOINED.
DR EMBL, M31192, AAA42261.1, JOINED.
DR EMBL, M31193, AAA42261.1, JOINED.
DR EMBL, M31194, AAA42261.1, JOINED.
DR EMBL, M31195, AAA42261.1, JOINED.
DR EMBL, M31196, AAA42261.1, JOINED.
DR EMBL, A19618, CAA01482.1, -;
DR PIR, A31597, A31597.
DR PIR, A35029, A35029.
DR HSSP, P00750, 1RTF.
DR MEROPS, S01.233, -;
DR InterPro, IPR001314, Chymotrypsin.

DR InterPro, IPR000561, EGF-like.
DR InterPro, IPR000083, Fibrinctn1.
DR InterPro, IPR000001, Kringle.
DR InterPro, IPR001254, Ser_protease_Try.
DR Pfam, PF00008, EGF, 1.
DR Pfam, PF00039, fn1, 1.
DR Pfam, PF00051, kringle, 2.
DR Pfam, PF00089, trypsin, 1.
DR PRINTS, PR00722, CHYMOTRYPSIN.
DR PRINTS, PR00018, KRINGLE.
DR ProDom, PD000395, Kringle, 2.
DR SMART, SM00181, EGF, 1.
DR SMART, SM00058, PNI, 1.
DR SMART, SM00130, KR, 2.
DR SMART, SM00020, Tryp_Spc, 1.
DR PROSITE, PS00022, EGF_1, 1.
DR PROSITE, PS01186, EGF_2, 1.
DR PROSITE, PS01253, FIBRONECTIN_1, 1.
DR PROSITE, PS00021, KRINGLE_1, 2.
DR PROSITE, PS00070, TRYPsin_DOM, 2.
DR PROSITE, PS00240, TRYPsin_DOM, 1.
DR PROSITE, PS00134, TRYPsin_HIS, 1.
DR PROSITE, PS00135, TRYPsin_SER, 1.
KW plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17
FT PROPEP 18 29
FT CHAIN 30 559
FT CHAIN 30 308
FT CHAIN 309 559
FT DOMAIN 36 78
FT DOMAIN 79 117
FT DOMAIN 124 205
FT DOMAIN 213 294
FT DOMAIN 309 559
FT ACT_SITE 355 355
FT ACT_SITE 404 404
FT ACT_SITE 510 510
FT ACT_SITE 510 510
FT DISULFID 38 68
FT DISULFID 66 75
FT DISULFID 83 94
FT DISULFID 88 105
FT DISULFID 107 116
FT DISULFID 124 205
FT DISULFID 145 187
FT DISULFID 176 200
FT DISULFID 213 294
FT DISULFID 234 276
FT DISULFID 265 289
FT DISULFID 297 428
FT DISULFID 340 356
FT DISULFID 348 417
FT DISULFID 442 516
FT DISULFID 474 490
FT DISULFID 506 534
FT CARBOHYD 149 149
FT CARBOHYD 481 481
FT CONFLICT 380 380
SQ SEQUENCE 559 AA; 62903 MM; 7DBD3809C1D1C921 CRC64;
Query Match 11.9%; Score 154; DB 1; Length 559;
Best local Similarity 32.4%; Pred. No. 1e-05;
Matches 45; Conservative 14; Mismatches 54; Indels 26; Gaps 6;
QY 4 CFWDNGHLVREDQTPAPGLRCLNMLDQSGLAGAPVS-----GAGNHSYCRUPDE 54
DB 124 CFEQGITRYRTQWSTAENGACIMW--NSSALSOKPI SARPPNAIKIGLGNHNTCRPPDR 181
QY 55 DFRGFWCVSGEAGVPEKRPCEIDLRCE-----TTSQALPAFTTEIOEASGPG 103
DB 182 DVK-FWCYVF-KAGKYTFEFCTPACPKGPTEDCYGVGKVTYGTSHSFTT--SKASGLPW 237

Qy 104 ADEVQVAPANALPARSEA 122
Db 238 NSMILIGKTYTAMPANSDA 256

RESULT 4
HGFA_MOUSE STANDARD; PRT; 653 AA.

AC Q9R098; Q9R098; 40, Created
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF activator) (HGFA).
GN HGFA_MOUSE (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RA Itoh H., Kataoka H., Kono H.;
RT "Mouse hepatocyte growth factor activator."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RA Yang J., Huan Y.S., Sehgal S., Kukes A., Brady C., Barasch J.,
RT "Activation of HGF by endogenous HGF activator is required for
RT metanephric kidney morphogenesis in vitro."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING
CC IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM (BY SIMILARITY).
CC - SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
CC DISULFIDE BOND (BY SIMILARITY).
CC - SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM (BY
CC SIMILARITY).
CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC - SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC - SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.
CC - SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.
CC - SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF039017; AAF02489.1; -
DR EMBL: AF224724; AAF34712.1; -
DR HSRP: P00763; IDPO.
DR MEROPS: S01.228; -
DR MGD: MGI:1859281; Hgfac.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF 2.
DR InterPro: IPR000562; FN Type II.
DR InterPro: IPR000083; Fibrinctn.
DR InterPro: IPR001254; Ser_protease_Ty.
DR Pfam: PF000039; EGF; 2.
DR Pfam: PF000039; fn1; 1.
DR Pfam: PF000440; fn2; 1.
DR Pfam: PF00051; kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00013; FNTYPEII.

DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 1.
DR ProDom: PD000995; FN_Type_II; 1.
DR SMART: SM00181; EGF 2.
DR SMART: SM00058; FN1; 1.
DR SMART: SM00059; FN2; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS00022; EGF 1; 2.
DR PROSITE: PS01186; EGF 2; 1.
DR PROSITE: PS01253; FIBRONECTIN_1; 1.
DR PROSITE: PS00023; FIBRONECTIN_2; 1.
DR PROSITE: PS00021; KRINGLE 1; 1.
DR PROSITE: PS00070; KRINGLE 2; 1.
DR PROSITE: PS00240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; 1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
KW Hydroxylase; glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 29
FT PROPEP 30 369
FT CHAIN 370 405
FT CHAIN 406 653
FT DOMAIN 105 145
FT DOMAIN 157 195
FT DOMAIN 197 237
FT DOMAIN 238 276
FT DOMAIN 283 364
FT DOMAIN 406 653
FT ACT_SITE 445 445
FT ACT_SITE 495 495
FT ACT_SITE 596 596
FT ACT_SITE 105 130
FT DISULFID 119 145
FT DISULFID 161 172
FT DISULFID 166 183
FT DISULFID 185 194
FT DISULFID 199 227
FT DISULFID 225 234
FT DISULFID 242 253
FT DISULFID 247 264
FT DISULFID 266 275
FT DISULFID 283 364
FT DISULFID 304 346
FT DISULFID 335 359
FT DISULFID 392 519
FT DISULFID 430 446
FT DISULFID 438 508
FT DISULFID 533 602
FT DISULFID 565 581
FT DISULFID 592 620
FT CARBOHYD 39 39
FT CARBOHYD 47 47
FT CARBOHYD 63 63
FT CARBOHYD 287 287
FT CARBOHYD 466 466
FT CARBOHYD 544 544
FT CONFLICT 164 164
SQ SEQUENCE 653 AA; 70567 MW; 88B4B2055DF7DC CRC64;

Query Match 11.6%; Score 150; DB 1; Length 653;
Best Local Similarity 32.6%; Pred. No. 2.7e-05;
Matches 46; Conservative 17; Mismatches 46; Indels 32; Gaps 7;

Qy 4 CPMNGHLTYREDTSPAPGLRCLNW-----LDAQSGLASAPVSGAGNHSYCRVPED 55
Db 283 CFLNGTETRYGVASTPAAGLSCLAMNSDLTYOELHVDV-VAAAVLLGLGPHAVCRPPDD 341

Qy 56 PRGPNCTVYSGENGAVKPCEDLACPETTSQALPAFTTEIQEASBEGADDEVQVAPANPA 115
Db 342 ER-PWCYVVKDNLALSWE-----YCRLLACESLARVHSQTFE-----TLA--A 380

QY 116 LPARSEMAVPGISQVR 116
DB 381 LP--ESAPAVRPTCGKHKR 399

RESULT 5
TPA_MOUSE STANDARD; PRT; 559 AA.

AC P11214;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
GN (t-PA) (t-plasminogen activator).

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88087303; PubMed=2826484;
RA Rickles R.J., Darrow A.L., Strickland S.;
RT "Molecular cloning of complementary DNA to mouse tissue plasminogen
RT activator mRNA and its expression during f9 teratocarcinoma cell
RT differentiation.";
RL J. Biol. Chem. 263:1563-1569(1988).

CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-308 CATALYZED BY PLASMIN. TISSUE KALLIKREIN OR FACTOR XA.
CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
CC or send an email to license@isb-slb.ch).

CC -----
CC EMBL, J03520, AAA0470.1; -
CC F1R, A29941; A29941.
CC HSSP, P00750; IASH.
CC MEROPS, S01.232; -
CC MGD, MGI:97610; Plat.
CC InterPro, IPR001314; Chymotrypsin.
CC InterPro, IPR000561; EGF-like
CC InterPro, IPR000083; Fibrinctn.
CC InterPro, IPR000001; Kringle.
CC InterPro, IPR001254; Ser_protease_Try.
CC Pfam, PF00008; EGF, 1.
CC Pfam, PF00039; fn1, 1.
CC Pfam, PF00051; kringle, 2.
CC Pfam, PF00089; trypsin, 1.
CC PRINTS, PR00722; CHYMOTRYPSIN.
CC PRINTS, PR00018; KRINGLE.
CC ProDom, PD000395; Kringle, 2.

DR SMART, SM00181; EGF, 1.
DR SMART, SM00058; FN1, 1.
DR SMART, SM00130; KR, 2.
DR SMART, SM00020; Tryp_Spc, 1.
DR PROSITE, PS00022; EGF_1, 1.
DR PROSITE, PS01186; EGF_2, 1.
DR PROSITE, PS01253; FIBRONECTIN_1, 1.
DR PROSITE, PS00021; KRINGLE_1, 2.
DR PROSITE, PS00707; KRINGLE_2, 2.
DR PROSITE, PS02407; TRYPSIN_DOM, 1.
DR PROSITE, PS00134; TRYPSIN_HIS, 1.
DR PROSITE, PS00135; TRYPSIN_SER, 1.
KW Plasminogen activator; Hydrolase; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17
FT PROPEP 18 29
FT CHAIN 30 559
FT CHAIN 30 508
FT CHAIN 309 559
FT DOMAIN 36 78
FT DOMAIN 79 117
FT DOMAIN 124 205
FT DOMAIN 213 294
FT DOMAIN 309 559
FT ACT SITE 355 355
FT ACT SITE 404 404
FT ACT SITE 510 510
FT ACT SITE 510 510
FT DISULFID 38 68
FT DISULFID 66 75
FT DISULFID 83 94
FT DISULFID 88 105
FT DISULFID 107 116
FT DISULFID 124 205
FT DISULFID 145 187
FT DISULFID 176 200
FT DISULFID 213 294
FT DISULFID 234 276
FT DISULFID 265 289
FT DISULFID 297 428
FT DISULFID 340 356
FT DISULFID 348 417
FT DISULFID 442 516
FT DISULFID 474 490
FT DISULFID 506 534
FT CARBOHYD 149 149
FT CARBOHYD 481 481
SQ SEQUENCE 559 AA; 63110 MW; 4ACEB57DCEA282A5 CRC64;

Query Match 11.4%; Score 147.5; DB 1; Length 559;
Best Local Similarity 37.0%; Pred. No. 3.7e-05;
Matches 34; Conservative 11; Mismatches 34; Indels 13; Gaps 4;

QY 4 CFMDNGHLYREBQSPARGCLNWLDAQGLASAPV-----GAGNHSYCRNPDE 54
DB 124 CFEEQITRGWSTYASGAECLNW--NSSVLSLKPYNRRPAIKLGLGNHYCNPDR 181

QY 55 DRGPWCYVSGAVPEKRPCEDLRCPEETSQ 86
DB 182 DLK-PWCYVF-KAGKYTFEFCSTPCKPKGSE 211

RESULT 6
HGFA_HUMAN STANDARD; PRT; 655 AA.
AC Q04756; Q14726;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
DE activator) (HGFA).
GN HGFA.

OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
ON NCI_TaxId=9606,
[1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC TISSUE=Liver, and Serum;
RK MEDLINE=932878; Pubmed=7683665;
RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,
RA Kitamura N.,
RT "Molecular cloning and sequence analysis of the cDNA for a human
RT serine protease responsible for activation of hepatocyte growth
RT factor. Structural similarity of the protease precursor to blood
RT coagulation factor XII".
RL J. Biol. Chem. 268:10024-10028(1993).
RN [2]
RP SEQUENCE OF 40-655 FROM N.A.
RA Zhao S., Odeh C.,
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY
CC CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM.
CC -1- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
CC DISULFIDE BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM.
CC -1- TISSUE SPECIFICITY: LIVER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
CC or send an email to license@1sb-sib.ch).

CC
DR EMBL; D14012; BAA03113.1; -
DR EMBL; Z69923; CAA93803.1; -
DR PIR; A46688; A46688.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.228; -
DR Genew; HGNC:4894; HGFAC.
DR MIM; 604552; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR000562; FN Type II.
DR InterPro; IPR000083; p18ctn1.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF 2.
DR Pfam; PF00039; Fn1; 1.
DR Pfam; PF00040; Fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; ENTYREIT.
DR PRINTS; PR00019; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR PRODOM; PD000995; FN_Type_II; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYPSIN; 1.
DR PROSITE; PS00022; EGF 1; 2.
DR PROSITE; PS01166; EGF 2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.

DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase, Glycoprotein, Plasma, Serine protease, Kringle, Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 30
FT PROPEP 31 372
FT CHAIN 373 407
FT CHAIN 408 655
FT DOMAIN 108 148
FT DOMAIN 160 198
FT DOMAIN 200 240
FT DOMAIN 241 279
FT DOMAIN 286 367
FT DOMAIN 408 655
FT ACT SITE 447 447
FT ACT SITE 497 497
FT ACT SITE 598 598
FT ACT SITE 108 133
FT DISULFID 122 148
FT DISULFID 164 175
FT DISULFID 169 186
FT DISULFID 188 197
FT DISULFID 202 230
FT DISULFID 228 237
FT DISULFID 245 256
FT DISULFID 250 267
FT DISULFID 269 278
FT DISULFID 266 367
FT DISULFID 307 349
FT DISULFID 338 362
FT DISULFID 394 521
FT DISULFID 432 448
FT DISULFID 440 510
FT DISULFID 535 604
FT DISULFID 567 583
FT DISULFID 594 622
FT CARBOHYD 48 48
FT CARBOHYD 290 290
FT CARBOHYD 468 468
FT CARBOHYD 492 492
FT CARBOHYD 546 546
FT CONFLICT 644 644
SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1B1862BD7 CRC64;
Query Match 11.4%; Score 146.5; DB 1; Length 655;
Best local Similarity 36.9%; Pred. No. 5,3e-05;
Matches 41; Conservative 10; Mismatches 41; Indels 19; Gaps 6;
Oy 4 CFMDNGHLTREDDTSPAPGLRLN-----IDAOSGLASAPVSGAGHSICRRPDD 55
Db 286 CFTGNGTGYGVAVSTASGLSCLANNSDLVQELHVDV-VGAALGLGHAVCRPNDD 344
Oy 56 PRGWCVCVSGEAGVP-----EKRPCEDLRCPEPTTSQALPAFTTEIOE-ASRG 101
Db 345 ER-PWCVVVDLSALSWCYRLACESTL---TRVQLSPDLATLTPBPASPG 390
RESULT 7
ID PLAN HUMAN STANDARD; PRT; 810 AA.
AC P00747;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OK NCBI_TaxId=9606;
 RN [11]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90202879; PubMed=2318848;
 RA Petersen T.E., Marzen M.R., Ichinose A., Davie E.W.;
 RT "Characterization of the gene for human plasminogen, a key proenzyme
 in the fibrinolytic system.";
 RL J. Biol. Chem. 265:6104-6111(1990).
 RN [12]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87162490; PubMed=1030813;
 RA Forsgren M., Røden B., Iversen M., Larsson K., Heden L.-O.;
 RT "Molecular cloning and characterization of a full-length cDNA clone
 for human plasminogen.";
 RL FEBS Lett. 213:254-260(1987).
 RN [13]
 RP SEQUENCE OF 20-810.
 RA Sottrup-Jensen L., Petersen T.E., Magnusson S.;
 RL Submitted (JUL-1977) to the PIR data bank.
 RN [14]
 RP SEQUENCE OF 292-810 FROM N.A.
 RX MEDLINE=85023311; PubMed=6148961;
 RA Malinowski D.P., Sadler J.E., Davie E.W.;
 RT "Characterization of a complementary deoxyribonucleic acid coding for
 human and bovine plasminogen.";
 RL Biochemistry 23:4243-4250(1984).
 RN [15]
 RP SEQUENCE OF 20-100.
 RX MEDLINE=75093329; PubMed=122932;
 RA Wiman B., Wallen P.;
 RT "Structural relationship between 'glutamic acid' and 'lysine' forms
 of human plasminogen and their interaction with the NH2-terminal
 activation peptide as studied by affinity chromatography.";
 RL Eur. J. Biochem. 50:489-494(1975).
 RN [16]
 RP SEQUENCE OF 95-580; 581-626; 657-700 AND 732-810.
 RA Sottrup-Jensen L., Claess H., Zajdel M., Petersen T.E., Magnusson S.;
 RL (In) Davidson J.F., Roman R.M., Samama M.M., Desnoyers P.C. (eds.);
 Progress in chemical fibrinolysis and thrombolysis, pp.3:191-209,
 Raven Press, New York (1978).
 RN [17]
 RP SEQUENCE OF 483-604.
 RX MEDLINE=76043692; PubMed=126663;
 RA Wiman B., Wallen P.;
 RT "Amino-acid sequence of the cyanogen-bromide fragment from human
 plasminogen that forms the linkage between the plasmin chains.";
 RL Eur. J. Biochem. 58:539-547(1975).
 RN [18]
 RP SEQUENCE OF 561-810.
 RX MEDLINE=77225245; PubMed=142009;
 RA Wiman B.;
 RT "Primary structure of the B-chain of human plasmin.";
 RL Eur. J. Biochem. 76:129-137(1977).
 RN [19]
 RP ACTIVE SITE.
 RX MEDLINE=73149248; PubMed=4694729;
 RA Robbins K.C., Bernabe P., Arrazdon L., Summaria L.;
 RT "Type primary structure of human plasminogen. II. The histidine loop
 of human plasmin: light (B) chain active center histidine sequence.";
 RL J. Biol. Chem. 248:1631-1633(1973).
 RN [10]
 RP ACTIVE SITE.
 RX MEDLINE=69234739; PubMed=4240117;
 RA Groskopf W.R., Summaria L., Robbins K.C.;
 RT "Studies on the active center of human plasmin. Partial amino acid
 sequence of a peptide containing the active center serine residue.";
 RL J. Biol. Chem. 244:3590-3597(1969).
 RN [11]
 RP OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.
 RX MEDLINE=8221905; PubMed=6919539;
 RA Trexler M., Vail Z., Patchy L.;

RT "Structure of the omega-aminocarboxylic acid-binding sites of human
 plasminogen. Arginine 70 and aspartic acid 56 are essential for
 binding of ligand by kringle 4.";
 RL J. Biol. Chem. 257:7401-7406(1982).
 RN [12]
 RP FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.
 RX MEDLINE=85054794; PubMed=6094526;
 RA Vail Z., Patchy L.;
 RT "The fibrin-binding site of human plasminogen. Arginines 32 and 34
 are essential for fibrin affinity of the kringle 1 domain.";
 RL J. Biol. Chem. 259:13690-13694(1984).
 RN [13]
 RP PHOSPHORYLATION SITE SER-597.
 RX MEDLINE=97345939; PubMed=9201958;
 RA Wang H., Pirook M., Brethauer R.K., Castellino F.J.;
 RT "Serine-578 is a major phosphorylation locus in human plasma
 plasminogen.";
 RL Biochemistry 36:8100-8106(1997).
 RN [14]
 RP CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=8818329; PubMed=3356193;
 RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamberling J.P.,
 Gerwig G.J., van Halbeek H., Vliegenhart J.F.;
 RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
 plasminogen. Species specificity in relation to stialylation and
 fucoylation patterns.";
 RL Eur. J. Biochem. 173:57-63(1988).
 RN [15]
 RP CARBOHYDRATE-LINKAGE SITE SER-268.
 RX MEDLINE=97207306; PubMed=9054441;
 RA Piate-Shepherd S.R., Stevens R.D., Andon N.L., Enghild J.J.,
 Pizzo S.V.;
 RT "Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of
 human plasminogen 2.";
 RL J. Biol. Chem. 272:7408-7411(1997).
 RN [16]
 RP CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
 RX MEDLINE=95042728; PubMed=7525077;
 RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
 Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
 RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
 suppression of metastases by a Lewis lung carcinoma.";
 RL Cell 79:315-328(1994).
 RN [17]
 RP CHARACTERIZATION OF ANGIOSTATIN.
 RX MEDLINE=9728710; PubMed=1657148;
 RA Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Madsen J.W.,
 Lapcevic R., Nacy C.A.;
 RT "A recombinant human angiostatin protein inhibits experimental primary
 cancer Res. 57:1329-1334(1997).
 RN [18]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.
 RX MEDLINE=92031502; PubMed=1657148;
 RA Mulichak A.M., Tulinsky A., Ravichandran K.G.;
 RT "Crystal and molecular structure of human plasminogen kringle 4
 refined at 1.9-A resolution.";
 RL Biochemistry 30:10576-10586(1991).
 RN [19]
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.
 RX MEDLINE=92031503; PubMed=1657149;
 RA Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;
 RT "The refined structure of the epsilon-aminocaproic acid complex of
 human plasminogen kringle 4.";
 RL Biochemistry 30:10589-10594(1991).
 RN [20]
 RP X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
 RA Sec B., Yamano A., Whitlow M., Teeter M.M.;
 RT "Structure of human plasminogen kringle 4 at 1.68 Angstrom and 277 K.
 A possible structural role of disordered residues.";
 RL Acta Crystallogr. D 53:169-178(1997).
 RN [21]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.

RX MEDLINE=96180681; PubMed=8611560;
 RA Mathews I.I., Vanderhoff-Hanaver P., Castellino F.J., Tulinsky A.;
 RT "Crystal structures of the recombinant kringle 1 domain of human
 RT plasminogen in complexes with the ligands epsilon-aminocaproic acid
 RT and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";
 RL Biochemistry 35:2567-2576(1996).
 RN [123]
 RN X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.
 RP MEDLINE=96198034; PubMed=9521645;
 RA Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,
 RA Castellino F.J.;
 RT "Structure and ligand binding determinants of the recombinant kringle
 RT 5 domain of human plasminogen.";
 RL Biochemistry 37:3258-3271(1998).
 RN [123]
 RP STRUCTURE BY NMR OF 96-184.
 RX MEDLINE=94237157; PubMed=8181475;
 RA Rejzante M.R., Llinas M.;
 RT "1H-NMR assignments and secondary structure of human plasminogen
 RT kringle 1.";
 RL Eur. J. Biochem. 221:927-937(1994).
 RN [124]
 RP STRUCTURE BY NMR OF 96-184.
 RX MEDLINE=94237158; PubMed=8181476;
 RA Rejzante M.R., Llinas M.;
 RT "Solution structure of the epsilon-aminohexanoic acid complex of
 RT human plasminogen kringle 1.";
 RL Eur. J. Biochem. 221:939-949(1994).
 RN [125]
 RP STRUCTURE BY NMR OF 183-354.
 RX MEDLINE=96194156; PubMed=8652577;
 RA Soehnlel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,
 RA Ricci E.B.;
 RT "Recombinant gene expression and 1H NMR characteristics of the
 RT kringle (2 + 3) supermodule: spectroscopic/functional individuality
 RT of plasminogen kringle domains.";
 RL Biochemistry 35:2357-2364(1996).
 RN [126]
 RP STRUCTURE BY NMR OF 374-461.
 RX MEDLINE=90219023; PubMed=2157850;
 RA Atkinson R.A., Williams R.J.P.;
 RT "Solution structure of the kringle 4 domain from human plasminogen by
 RT 1H nuclear magnetic resonance spectroscopy and distance geometry.";
 RL J. Mol. Biol. 212:541-552(1990).
 RN [127]
 RP VARIANTS PHE-374 AND THR-620.
 QY 4 CFWDNGHYREDQSPAPGLRCUMLDA-----QSGLASAPVSGAGNHCYCRNPDDPFG 58
 DB 103 CTTGNGKRYRGTMTKNGITCKKMSSTSPHRRPRPAHPEGL-ENNYCRNPDDPFG 161
 QY 59 PWCYVSGEAGVPEKR--PCEDLRCP 82
 DB 162 PWCYTTD---PEKRYDCILCEB 183
 RESULT 8
 PLAN RAT STANDARD; PRT; 169 AA.
 AC 001177;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Plasminogen (EC 3.4.21.7) (Fragment).
 GN Plg.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OC NCBI_TaxID=10116;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=91250378; PubMed=1645711;
 RA Karalis J.J., Macker S.P.;
 RT Identification of the rat Heymann nephritis autoantigen (GP330) as a
 RT receptor site for plasminogen.";
 RL J. Biol. Chem. 266:10825-10829(1991).
 CC -1- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
 CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
 CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
 CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
 CC GRANULOSA FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
 CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
 CC AS C1 AND C5. IT CLEAVES FIBRIN FIBRONECTIN, THROMBOSPONDIN,
 CC LAMININ AND VON WILLEBRAND FACTOR.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
 CC higher selectivity than trypsin. Converts fibrin into soluble
 CC products.
 CC -1- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
 CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
 CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
 CC -1- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
 CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; M62832; AAA1884.1; -
 CC FIR; A40522; A40522.
 CC DR HSSE; P00747; IPMK.
 CC DR MEROP6; S01.233; -
 CC DR InterPro; IPR000001; Kringle.
 CC DR InterPro; IPR001254; Ser protease_Try.
 CC DR Pfam; PF00051; Kringle; 2.
 CC DR ProDom; PD000395; Kringle; 2.
 CC DR SMART; SMO0130; KR; 1.
 CC DR PROSITE; PS00021; KRINGLE_1; 1.
 CC DR PROSITE; PS00070; KRINGLE_2; 2.
 CC DR PROSITE; PS50240; TRYPsin_DOM; PARTIAL.
 CC DR PROSITE; PS50134; TRYPsin_HIS; PARTIAL.
 CC DR PROSITE; PS00135; TRYPsin_SER; PARTIAL.
 CC KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
 CC KW Tissue remodeling; Blood coagulation; Kringle; Repeat.
 CC FT NON TER 1 1
 CC FT DOMAIN <1 10 KRINGLE 3 (BY SIMILARITY).
 CC FT DOMAIN 34 112 KRINGLE 4 (BY SIMILARITY).
 CC FT DOMAIN 139 >169 KRINGLE 5 (BY SIMILARITY).
 CC FT DISULFD 34 112 BY SIMILARITY.
 CC FT DISULFD 55 95 BY SIMILARITY.
 CC FT DISULFD 83 107 BY SIMILARITY.
 CC FT NON TER 169 169
 CC SQ SEQUENCE 169 AA; 18401 MW; 77A54214C49D010C CRC64;
 Query Match 10.7%; Score 138; DB 1; Length 169;
 Best Local Similarity 31.8%; Pred. No. 6.3e-05;
 Matches 35; Conservative 13; Mismatches 48; Indels 14; Gaps 5;
 QY 4 CFWDNGHYREDQSPAPGLRCUMLDA-----QSGLASAPVSGAGNHCYCRNPDDPFG 58
 DB 34 CTTGNGKRYRGTSTTYNGKCKCQSWSMTPHSHSKTPANFPDGG-ENNYCRNPDDPFG 92
 QY 59 PWCYVSGEAGVPEKR--PCEDLRCPETTSQALPAFTTEIGQASGCPGAD 106
 DB 93 PWCFTTD---PSVMEYCNMKRCSEFGGV--AESAIIVPVPAFGTSE 136

| | | | | |
|------------------|--|-----------|------|---------------------------------------|
| | RESULT | 9 | | |
| ID | UROK_PAPCY | STANDARD; | PRT; | 433 AA. |
| AC | p16277; | | | |
| DT | 01-APR-1990 (Rel. 14, Created) | | | |
| DT | 01-APR-1990 (Rel. 14, Last sequence update) | | | |
| DT | 15-JUN-2002 (Rel. 41, Last annotation update) | | | |
| DE | Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (unp) | | | |
| GN | Papio. | | | |
| NCBI_TaxID=9556; | | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; | | | |
| OC | Cercopithecinae; Papio. | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TITISURE-phoracic aorta; | | | |
| RX | MEDLINE#90287734, PubMed#2113276; | | | |
| RA | A J.P.T., Wang I.W., Glover A.W.; | | | |
| RT | Nucleotide and deduced amino acid sequences of baboon urokinase-type | | | |
| RT | plasminogen activator." | | | |
| RL | Nucleic Acids Res. 18:3411-3411(1990). | | | |
| CC | -1- CATALYTIC ACTIVITY: Specific cleavage of Arg- -Val bond in | | | |
| CC | plasmaogen to form plasmin. | | | |
| CC | -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS | | | |
| CC | OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A | | | |
| CC | LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW | | | |
| CC | MOLECULAR MASS FORM TO YIELD A SHORT A CHAIN (BY SIMILARITY). | | | |
| CC | -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. | | | |
| CC | -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN. | | | |
| CC | -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN. | | | |
| CC | ----- | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL Outstation - | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on ways | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | |
| CC | entities requires a license agreement (See http://www.isb-stb.ch/announce/ | | | |
| CC | or send an email to license@isb-stb.ch). | | | |
| CC | ----- | | | |
| DR | EMBL; XS1935; CAAB6200.1; - | | | |
| DR | PIR; S14687; URGANY. | | | |
| DR | HSSP; P00749; 11MM. | | | |
| DR | MEROBS; S01231; -- | | | |
| DR | InterPro; IPR001314; Chymotrypsin. | | | |
| DR | InterPro; IPR000561; EGF-like. | | | |
| DR | InterPro; IPR000001; Kringle. | | | |
| DR | InterPro; IPR001254; Ser protease_Try. | | | |
| DR | Pfam; PF000053; kringle_1. | | | |
| DR | Pfam; PF00089; tryptsin_1. | | | |
| DR | PRINTS; PR00722; CHYMOTRYPSIN. | | | |
| DR | ProDom; PD000395; Kringle; 1. | | | |
| DR | SMART; SM00181; EGF_1. | | | |
| DR | SMART; SM00130; KR; 1. | | | |
| DR | SMART; SM00020; TRYPSIN_SPC; 1. | | | |
| DR | PROSITE; PS00022; EGF_1; 1. | | | |
| DR | PROSITE; PS01186; EGF_2; FALSE NEG. | | | |
| DR | PROSITE; PS00021; KRINGLE_1; 1. | | | |
| DR | PROSITE; PS50070; KRINGLE_2; 1. | | | |
| DR | PROSITE; PS50240; TRYPsin_DOM; 1. | | | |
| DR | PROSITE; PS00134; TRYPsin_HIS; 1. | | | |
| DR | PROSITE; PS00135; TRYPsin_SER; 1. | | | |
| KW | Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein; | | | |
| KM | Kringle; EGF-like domain; Zymogen; Signal. | | | |
| FT | SIGNAL | 1 | 20 | POTENTIAL. |
| FT | CHAIN | 1 | 433 | UROKINASE-TYPE PLASMINOGEN ACTIVATOR. |
| FT | CHAIN | 21 | 176 | CHAIN A (BY SIMILARITY). |
| FT | CHAIN | 155 | 176 | SHORT A CHAIN (A1) (BY SIMILARITY). |
| FT | CHAIN | 178 | 433 | CHAIN B (BY SIMILARITY). |
| FT | DOMAIN | 26 | 62 | EGF-LIKE. |
| FT | DOMAIN | 66 | 150 | KRINGLE. |

[illegible]

DR MEROS; S01.231; -.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Ser protease_Try.
 DR Pfam: PF00051; kringle; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR ProDom: PD000395; Kringle; 1.
 DR SMART: SM00130; KR; 1.
 DR SMART: SM00020; TRYP_SPE; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS0186; EGF_2; FALSE_NEG.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS00070; KRINGLE_2; 1.
 DR PROSITE: PS00240; TRYPsin_DOM; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; 1.
 DR PROSITE: PS00135; TRYPsin_SER; 1.
 DR Plasmidogen activation; Hydrolase; Serine protease; Glycoprotein;
 Kringle; EGF-like domain; Zymogen; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 21 188 CHAIN A (BY SIMILARITY).
 FT DOMAIN 190 442 CHAIN B (BY SIMILARITY).
 FT DOMAIN 29 65 EGF-LIKE.
 FT DOMAIN 72 153 KRINGLE.
 FT DOMAIN 154 189 CONNECTING PEPTIDE.
 FT DOMAIN 190 442 SERINE PROTEASE.
 FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .).
 FT DISULFID 33 41 BY SIMILARITY.
 FT DISULFID 35 53 BY SIMILARITY.
 FT DISULFID 55 64 BY SIMILARITY.
 FT DISULFID 179 310 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 220 236 BY SIMILARITY.
 FT DISULFID 228 299 BY SIMILARITY.
 FT DISULFID 324 393 BY SIMILARITY.
 FT DISULFID 356 372 BY SIMILARITY.
 FT DISULFID 383 411 BY SIMILARITY.
 FT ACT_SITE 235 235 CHARGE RELAY SYSTEM.
 FT ACT_SITE 286 286 CHARGE RELAY SYSTEM.
 FT ACT_SITE 387 387 CHARGE RELAY SYSTEM.
 FT CONFLICT 241 241 O -> H (IN REF. 1; CAA25806).
 FT CONFLICT 242 242 Q -> H (IN REF. 1; CAA25511).
 FT CONFLICT 288 288 A -> GS (IN REF. 1; CAA25806).
 SQ SEQUENCE 442 AA; 4916 MM; E82FCEFS0132IEE CRC64;

Query Match 10.5%; Score 135.5; DB 1; Length 442;
 Best Local Similarity 36.9%; Pred. No. 0.0003;
 Matches 38; Conservative 7; Mismatches 37; Indels 21; Gaps 5;

OY 4 CFWNGHLYREDTSPAPGLRCINWLDAGSG---ASAPVS---GAGNHSYCRNPEDP 56
 DB 72 CFEENGHSYRKANTNTNGRPRCLPMSATVLTNTYAHKPDALQGLGKNYKCNPPD-NQ 130
 OY 57 RGPWCYVS-----GEAYP-----EKRPCEDLRCPTTSQ 86
 DB 131 RRPWCYVVGKLGKLVQECWVNCSGESHRPVDGKNPSTBE 173

RESULT 11
 URRB DESRO
 ID URRB DESRO STANDARD; PRT; 431 AA.
 AC P98121.
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA beta).
 OS Desmodus rotundus (Vampire bat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 OC Desmodontinae; Desmodus.
 OX NCBI_TaxID=9430;

EN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=92039036; PubMed=1937019;
 RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
 RA Alagon A., Donner P., Schleuning W.D.;
 RT "The plasminogen activator family from the salivary gland of the
 RT vampire bat Desmodus rotundus: cloning and expression.";
 RL Gene 105:229-237(1991).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=93393059; PubMed=1309059;
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
 RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Wilt W.,
 RA Donner P.;
 RT "Plasminogen activators from the saliva of Desmodus rotundus (common
 RT vampire bat): unique fibrin specificity."
 RL Ann. N.Y. Acad. Sci. 667:385-403(1992).
 CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
 CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
 CC AGENT.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTRAINS 1 KRINGLE DOMAIN.
 CC -1- SIMILARITY: CONTRAINS 1 EGF-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M63989; AAA31594.1; -.
 DR HSSP: P98119; 1A51.
 DR MEROPS; S01.239; -.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Ser protease_Try.
 DR Pfam: PF00089; EGF; 1.
 DR Pfam: PF00051; kringle; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR ProDom: PD000395; Kringle; 1.
 DR SMART: SM00181; EGF; 1.
 DR SMART: SM00130; KR; 1.
 DR SMART: SM00020; TRYP_SPE; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS0186; EGF_2; 1.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS00070; KRINGLE_2; 1.
 DR PROSITE: PS00240; TRYPsin_DOM; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; 1.
 DR PROSITE: PS00135; TRYPsin_SER; 1.
 KW Plasmidogen activation; Hydrolase; Serine protease; Glycoprotein;
 KW Kringle; EGF-like domain; Signal; Multigene family.
 FT SIGNAL 1 36
 FT CHAIN 37 431 POTENTIAL.
 FT DOMAIN 37 75 SALIVARY PLASMINOGEN ACTIVATOR BETA.
 FT DOMAIN 82 163 EGF-LIKE.
 FT DOMAIN 179 431 KRINGLE.
 FT ACT_SITE 226 226 SERINE PROTEASE.
 FT ACT_SITE 275 275 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 382 382 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 41 52 BY SIMILARITY.
 FT DISULFID 46 63 BY SIMILARITY.
 FT DISULFID 65 74 BY SIMILARITY.
 FT DISULFID 82 163 BY SIMILARITY.

FT DISULFID 103 145 BY SIMILARITY.
 FT DISULFID 134 158 BY SIMILARITY.
 FT DISULFID 168 299 BY SIMILARITY.
 FT DISULFID 211 227 BY SIMILARITY.
 FT DISULFID 219 288 BY SIMILARITY.
 FT DISULFID 313 388 BY SIMILARITY.
 FT DISULFID 345 361 BY SIMILARITY.
 FT DISULFID 378 406 BY SIMILARITY.
 FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 431 AA; 48221 MW; 6995E675B162CBF CRC64;

Query Match 10.4%; Score 134; DB 1; Length 431;
 Best Local Similarity 36.6%; Pred. No. 0.00039;
 Matches 27; Conservative 9; Mismatches 20; Indels 14; Gaps 3;
 QY 4 CPMNGHLVREDQTSPPAFGLRCNLMDAOSGL-----ASAPVSGAGNHSYCRNPD 53
 DB 82 CYKDGVTYRGWSTSESGACINW---NSNLTTRTYNGRSDDATLTGLGNHNYCRNPD 136
 QY 54 EDRGPMCYV 63
 DB 139 NNSK-PMCTV 147

RESULT 12
 ID _UR22 DESRO STANDARD; PRT; 477 AA.
 AC P15638;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSPA
 alpha-2) (Bat-PA) (t-plasminogen activator).
 OS Desmodus rotundus (Vampire bat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 OC Desmodontinae; Desmodus.
 NCBI_TaxID=9430;
 RN [1]
 RP SEQUENCE FROM N. A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=92039036; PubMed=1937019;
 RA Kraetzschmar J., Haendler B., Langer G., Boisdol W., Bringmann P.,
 RA Alagon A., Donner P., Schleuning W.D.,
 RT "The plasminogen activator family from the salivary gland of the
 RT vampire bat Desmodus rotundus: cloning and expression.";
 RL Gene 105:229-237(1991).
 RN [2]
 RP SEQUENCE FROM N. A. AND PARTIAL SEQUENCE.
 RC TISSUE=Salivary gland;
 RX MEDLINE=90036867; PubMed=2509450;
 RA Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,
 RA Regester R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.,
 RT "Isolation, characterization, and cDNA cloning of a vampire bat
 RT salivary plasminogen activator.";
 RL J. Biol. Chem. 264:17947-17952(1989).
 RN [3]
 RP CHARACTERIZATION.
 RC MEDLINE=93393059; PubMed=1309059;
 RA Schleuning W.-D., Alagon A., Boisdol W., Bringmann P., Petri T.,
 RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
 RA Donner P.,
 RT "Plasminogen activators from the saliva of Desmodus rotundus (common
 RT vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
 CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
 CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
 CC AGENT.
 CC -1- CATALYTIC ACTIVITY. Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN
 CC THE PRESENCE OF FIBRIN I.

CC -1- SUBUNIT: MONOMER.
 CC -1- DOMAIN: THE FIBRONECTIN TYPE-1 DOMAIN MEDIATES BINDING TO FIBRIN,
 CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
 CC STIMULATION OF ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE 1 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M63988; AAA31593.1; -;
 CC EMBL; J05082; AAA31596.1; -;
 CC PIR; A34369; A34369.
 CC HSP; P98119; 1A51.
 CC MEROS; S01232; -;
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR000561; EGF-like.
 CC InterPro; IPR000083; Fibrinctn1.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00008; EGF_1.
 CC Pfam; PF00039; fn1; 1.
 CC Pfam; PF00051; kringle; 1.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC ProDom; PD000395; kringle; 1.
 CC SMART; SM00181; EGF; 1.
 CC SMART; SM00130; KR; 1.
 CC SMART; SM00020; Tryp_Spc; 1.
 CC PROSITE; PS00022; EGF_1; 1.
 CC PROSITE; PS01186; EGF_2; 1.
 CC PROSITE; PS01253; FIBRONECTIN_1; 1.
 CC PROSITE; PS00021; KRINGLE_1; 1.
 CC PROSITE; PS50070; KRINGLE_2; 1.
 CC PROSITE; PS50040; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC K1 Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
 CC Kringle; EGF-like domain; Signal; Multigene family.
 CC SIGNAL
 CC FT CHAIN 1 36
 CC FT CHAIN 37 477
 CC FT DOMAIN 40 82
 CC FT DOMAIN 83 121
 CC FT DOMAIN 128 209
 CC FT DOMAIN 225 477
 CC FT ACT_SITE 321 321
 CC FT ACT_SITE 428 428
 CC FT DISULFID 42 72
 CC FT DISULFID 70 79
 CC FT DISULFID 87 98
 CC FT DISULFID 92 109
 CC FT DISULFID 111 120
 CC FT DISULFID 128 209
 CC FT DISULFID 149 191
 CC FT DISULFID 180 204
 CC FT DISULFID 214 345
 CC FT DISULFID 257 273
 CC FT DISULFID 265 334
 CC FT DISULFID 359 434
 CC FT DISULFID 391 407
 CC FT DISULFID 424 452
 CC FT DISULFID 424 452
 CC FT CARBOHYD 185 185
 CC FT CARBOHYD 398 398
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).


```
FT CONFLICT 403 403 N -> K (IN REF. 2).
FT CONFLICT 417 417 Y -> H (IN REF. 2).
FT CONFLICT 435 435 M -> R (IN REF. 2).
SQ SEQUENCE 477 AA; 53719 MW; 17486555C0E5077C CRC64;

Query Match 10.4%; Score 134; DB 1; Length 477;
Best Local Similarity 38.6%; Pred. No. 0.00044;
Matches 27; Conservative 9; Mismatches 20; Indels 14; Gaps 3;

OY 4 CPWDGHLHYREQPSAPGLRGLTMDLADQSGI-----ASAPVSGAGNHSYCRNPD 53
Db 128 CKXDGVTYRGWSTSEGAQCTNM---NSNLTTRTYNGRRSDAITYLGLGNHNTCRNPD 184
OY 54 BDPGRPCYV 63
Db 185 NNSK-PWCYV 193

RESULT 13
UROK_HUMAN STANDARD; PRT; 431 AA.
ID _UROK_HUMAN 015844; Q16618; Q969W6;
AC 21-JUL-1986 (Rel. 01, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
DE (U-plasminogen activator).
GN PLAU.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxId=9606;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=85215647; PubMed=2987867;
RA Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Biasi F.;
RT "The human urokinase-plasminogen activator gene and its promoter.";
RL Nucleic Acids Res. 13:2759-2771(1985).
RN 12
RP SEQUENCE FROM N.A.
RA Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,
RA Steffens G.J., Heynaker H.U.;
RT "Cloning and expression of the gene for pro-urokinase in Escherichia
RT coli.";
RL Biotechnology 3:923-929(1985).
RN 13
RP SEQUENCE FROM N.A.
RX MEDLINE=86056954; PubMed=2415429;
RA Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Arimura H.,
RA Nishida M., Suyama T.;
RT "Molecular cloning of cDNA coding for human preprourokinase.";
RL Gene 36:183-188(1985).
RN 14
RP SEQUENCE FROM N.A.
RX MEDLINE=85203359; PubMed=3888571;
RA Jacobs P., Cravador A., Lortiau R., Brockly F., Coliau B., Chuchana P.,
RA van Elsen A., Herzog A., Bollen A.;
RT "Molecular cloning, sequencing, and expression in Escherichia coli of
RT human preprourokinase cDNA.";
RL DNA 4:139-146(1985).
RN 15
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN 16
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strusberg R.;
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN 17
RP SEQUENCE OF 66-431 FROM N.A.
RX MEDLINE=84272706; PubMed=6589620;
```

```
RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Biasi F.;
RT "Identification and primary sequence of an unspliced human urokinase
RT poly(A)+ RNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
RN 18
RP SEQUENCE OF 21-177.
RX MEDLINE=83055084; PubMed=6754569;
RA Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,
RA Flohe L.;
RT "The primary structure of high molecular mass urokinase from human
RT urine. The complete amino acid sequence of the A chain.";
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
RN 19
RP SEQUENCE OF 156-176 AND 179-224.
RX MEDLINE=83003608; PubMed=6749491;
RA Schaller J., Nick H., Rickl E.E., Gilleesen D., Lergier W.,
RA Struder R.O.;
RT "Human low-molecular-weight urinary urokinase. Partial
RT characterization and preliminary sequence data of the two polypeptide
RT chains.";
RL Eur. J. Biochem. 125:251-257(1982).
RN 110
RP SEQUENCE OF 158-410.
RX MEDLINE=83055099; PubMed=6754572;
RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.;
RT "The complete amino acid sequence of low molecular mass urokinase
RT from human urine.";
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
RN 111
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=96000858; PubMed=8591045;
RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
RA Dobson C.M., Stuart D.I., Jones E.Y.;
RT "The crystal structure of the catalytic domain of human
RT urokinase-type plasminogen activator.";
RL Structure 3:681-691(1995).
RN 112
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
RX MEDLINE=20266327; PubMed=10805774;
RA Speil S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,
RA Bode W., Magdolen V., Huber R., Moroder L.;
RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly
RT selective inhibitors of human urokinase.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
RN 113
RP STRUCTURE BY NMR.
RX MEDLINE=89127526; PubMed=2536903;
RA Oswald R.E., Bogusky M.U., Bamberger M., Smith R.A.G., Dobson C.M.;
RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
RT dimensional NMR.";
RL Nature 337:579-582(1989).
RN 114
RP STRUCTURE BY NMR OF 67-155.
RX MEDLINE=93003110; PubMed=1327118;
RA Li X., Smith R.A.G., Dobson C.M.;
RT "Sequential 1H NMR assignments and secondary structure of the kringe
RT domain from urokinase.";
RL Biochemistry 31:9562-9571(1992).
RN 115
RP STRUCTURE BY NMR OF 67-155.
RX MEDLINE=94149701; PubMed=8107091;
RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.;
RT "Solution structure of the kringe domain from urokinase-type
RT plasminogen activator.";
RL J. Mol. Biol. 235:1548-1559(1994).
RN 116
RP VARIANT LEU-141.
RX MEDLINE=96186279; PubMed=8652631;
RA Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takehashi K.,
RA Sawasaki Y., Hanada K.;
RT "Characterization of single chain urokinase-type plasminogen
RT activator with a novel amino-acid substitution in the kringe
RT structure.";
```

RL Biochim. Biophys. Acta 1293:83-89(1996).
 RN [17]
 RP VARIANT LEU-141.
 RX MEDLINE=97218551; PubMed=9065988;
 RA Come B., Berczy M., Belin D.;
 RT "Detection of polymorphisms in the human urokinase-type plasminogen
 RT activator gene";
 RL Thromb. Haemost. 77:434-435(1997).
 RN [16]
 RP ERRATUM.
 RA Come B., Berczy M., Belin D.;
 RL Thromb. Haemost. 78:973-973(1997).
 RN [19]
 RP VARIANT LEU-141.
 RX MEDLINE=97337920; PubMed=9194591;
 RA Turkmen B., Schmitt M., Schmalzfeldt B., Trommler P., Hell W.,
 RA Teutzbürg S., Graeff H., Magdolen V.;
 RT "Mutational analysis of the genes encoding urokinase-type plasminogen
 RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer";
 RL Electrophoresis 18:686-689(1997).
 CC -1- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR
 CC THERAPY OF THROMBOLYTIC DISORDERS.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
 CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
 CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.
 CC -1- PHARMACEUTICAL: Available under the name Abbotkinase (Abbott). Used
 CC in Pulmonary Embolism (PE) to initiate fibrinolysis.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, X02419; CAA26268.1; -;
 DR EMBL, M15476; AAA61253.1; -;
 DR EMBL, D00244; BAA00175.1; -;
 DR EMBL, D1143; BAA01919.1; -;
 DR EMBL, X02760; CAA26535.1; -;
 DR EMBL, AF377330; AAK53822.1; -;
 DR EMBL, BC013575; AAK13575.1; -;
 DR EMBL, K03226; AAC97138.1; -;
 DR EMBL, K02286; AAA61252.1; -;
 DR EMBL, A21571; CAA01559.1; -;
 DR EMBL, A18397; CAA01390.1; -;
 DR PIR, A00931; UKHU.
 DR PIR, A32974; A32974.
 DR PDB, 1KDU; 31-OCT-93.
 DR PDB, 1LMW; 29-JAN-96.
 DR PDB, 1URK; 08-MAY-95.
 DR PDB, 1EON; 17-MAY-00.
 DR MEROPS, S01.231; -;
 DR GlycoSuiteDB, P00749; -;
 DR Genew, HGNC:9052; PLAU.
 DR MIM, 191840; -;
 DR InterPro, IPR001314; Chymotrypsin.
 DR InterPro, IPR000561; EGF-like.
 DR InterPro, IPR000001; Kringle.
 DR InterPro, IPR01254; Ser_protease_Try.
 DR Pfam, PF00051; kringle; 1.
 DR Pfam, PF00089; trypsin; 1.
 DR PRINTS, PR00722; CHYMOTRYPSIN.
 DR PRODOM, PD000395; Kringle; 1.
 DR SMART, SM00181; EGF, 1.
 DR SMART, SM00130; KR; 1.

DR SMART, SM00020; Tryp_Spc; 1.
 Query Match 10.2%; Score 132; DB 1; Length 431;
 Best Local Similarity 32.7%; Pred. No. 0.00058;
 Matches 33; Conservative 13; Mismatches 33; Indels 22; Gaps 5;
 Qy 4 CFMDNGHLYREDQTSPPAGRLCNWLD-----OSGLASAPVSGAGNHSYCRNPDEDP 56
 Db 70 CYEGNGHFYRGKASTDTMGRCPLPMSATVYLOQTYAHNSDQLQLGNGYCRNPD-NR 128
 Qy 57 RGFMCYVSGEAGVPEKRP-----CEDLRCPETTSQAL 88
 Db 129 RRPWCYV--QVGL--KPLVQECVHDCADGKKSPSPPEEL 164
 RESULT 14
 HGFL_MOUSE STANDARD, PRT; 716 AA.
 AC P26928;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hepatocyte growth factor-like protein precursor (Macrophage
 DE stimulatory protein) (MSP).
 GN MST1 OR HGFL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C; TISSUE=Liver;
 RX MEDLINE=92002017; PubMed=1832957;
 RA Friesner Degen S.O., Stuart L.A., Han S., Jamison C.S.;
 RT "Characterization of the mouse cDNA and gene coding for a hepatocyte
 RT growth factor-like protein: expression during development.";
 RL Biochemistry 30:9781-9791(1991).
 CC -1- FUNCTION: PROBABLY HAS NO PROTEOLYTIC ACTIVITY, SINCE CRUCIAL AA
 CC CHARACTERISTIC OF SERINE PROTEASES CATALYTIC SITES ARE NOT
 CC CONSERVED.
 CC -1- TISSUE SPECIFICITY: LIVER. LOWER LEVELS IN LUNG, PLACENTA AND
 CC ADRENAL.
 CC -1- DEVELOPMENTAL STAGE: IS EXPRESSED AT LOW LEVELS DURING GESTATION.
 CC JUST BEFORE BIRTH THE LEVEL INCREASES DRAMATICALLY AND REMAINS
 CC STABLE AFTERWARDS.
 CC -1- PTM: MAY BE CLEAVED AFTER AA 488, TO YIELD A TWO-CHAIN MOLECULE
 CC HELD TOGETHER BY DISULFIDE BONDS, OR CLEAVED INTO TWO SEPARATE
 CC POLYPEPTIDES.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, M74180; AAA50166.1; -;
 DR EMBL, M74181; AAA50167.1; -;
 DR HSSP, P00747; IKRN.
 DR MEROPS, S01.975; -;
 DR MGD, MGI:96080; Hgfl.
 DR InterPro, IPR001314; Chymotrypsin.
 DR InterPro, IPR000001; Kringle.
 DR InterPro, IPR003014; PAN.
 DR InterPro, IPR003609; Pan_app.
 DR InterPro, IPR01254; Ser_protease_Try.
 DR Pfam, PF00024; PAN; 1.
 DR Pfam, PF00051; kringle; 4.
 DR Pfam, PF00089; trypsin; 1.
 DR PRINTS, PR00722; CHYMOTRYPSIN.

```

DR PRINTS; PR00018; KRINGLE.
DR Prodom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AB; 1.
DR SMART; SM00020; TRYP_SPE; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS50070; KRINGLE_2; 4.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
KW Kringle; Glycoprotein; Serine protease homolog; Repeat; Signal.
FT SIGNAL 1
FT CHAIN 1
FT CHAIN 19
FT DOMAIN 19
FT DOMAIN 110
FT DOMAIN 191
FT DOMAIN 292
FT DOMAIN 379
FT DOMAIN 457
FT DOMAIN 489
FT DISULFID 56
FT DISULFID 60
FT DISULFID 110
FT DISULFID 131
FT DISULFID 157
FT DISULFID 191
FT DISULFID 194
FT DISULFID 212
FT DISULFID 240
FT DISULFID 292
FT DISULFID 313
FT DISULFID 341
FT DISULFID 379
FT DISULFID 400
FT DISULFID 428
FT DISULFID 477
FT DISULFID 512
FT DISULFID 528
FT DISULFID 607
FT DISULFID 637
FT DISULFID 662
FT CARBOHYD 72
FT CARBOHYD 173
FT CARBOHYD 305
FT CARBOHYD 620
FT CONFLICT 19
SQ SEQUENCE 716 AA; 80588 MW; BBCE02EF85213ACC CRC64;

Query Match 10.1%; Score 130.5; DB 1; Length 716;
Best Local Similarity 30.7%; Pred. No. 0.0014;
Matches 35; Conservative 7; Mismatches 37; Indels 35; Gaps 5;

OY 4 CFWDNGHLYREDQTSAPGLRCLINW---LDAQSGLASAFVSGAGNHSYCRNDEDPGRGW 60
DB 110 CIMDNGVSTRGVARTAGGLPCQAMRRFPNDHKYTPPKNGI-BENFRNPDGDPRGPW 168

OY 61 CYVS-----GEAGVPEK-RPCE--DIRCET 83
DB 169 CYTTHNSVRFQSGIKTCREAVCVLCNGEDYRGEVDVTEGREGCORWDLQHPHS 222

RESULT 15
PLMN_ERIEU STANDARD; PRT; 810 AA.
AC 029485;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7).
GN PLG.
OS Eriaceae eurypaeus (Western European hedgehog).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Insectivora; Eriaceidae; Eriaceinae; Eriaceus.
OX NCBI_TaxID=9365;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;

```

```

RX MEDLINE=96025778; PubMed=7592597;
RA Lawn R.M., Boommark N.W., Schwartz K., Lindahl G.E., Wade D.P.,
RA Byrne C.D., Fong K.Y., Meer K., Patchy L.;
RT "The recurring evolution of lipoprotein(a). Insights from cloning of
RT hedgenet apolipoprotein(a).";
RL U. Biol. Chem. 270:24004-24009(1995).
RN [2]
RP REVISIONS.
RA Lawn R.M.;
Submitted (Feb-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC GRAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC ACTIVATOR, COLLAGENASE AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRINECTIN, THROMBOSPONDIN,
CC LAMININ AND VON WILLEBRAND FACTOR.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -1- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC ACTIVATORS. BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC -1- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL; U31171; AAC48717.1; -.
DR HSSP; P00747; 1PMK.
DR MEROPS; S01.233; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; PAN_app.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PR00024; PAN; 1.
DR Pfam; PR00051; kringle; 5.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR Prodom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 5.
DR SMART; SM00473; PAN_AB; 1.
DR SMART; SM00020; TRYP_SPE; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS50070; KRINGLE_2; 5.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
KW Signal.
FT SIGNAL 1
FT CHAIN 1
FT CHAIN 20
FT CHAIN 583
FT CHAIN 810
FT DOMAIN 103
FT DOMAIN 181
FT DOMAIN 185
FT DOMAIN 275
FT DOMAIN 379
FT DOMAIN 482
FT ACT_SITE 622
FT ACT_SITE 622

```

FT ACT_SITE 665 665 CHARGE RELAY SYSTEM.
 FT ACT_SITE 760 760 CHARGE RELAY SYSTEM.
 FT CAREQHYD 339 339 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 810 AA; 90902 MM; 8E75780946017A16 CRC64;

Query Match 10.0%; Score 129; DB 1; Length 810;
 Best Local Similarity 34.1%; Pred. No. 0.0021;
 Matches 31; Conservative 7; Mismatches 31; Indels 22; Gaps 4;

Qy 4 CFWDNGHLYREDQTSBAPGLRCLNWLDAOSGLASAP-----VSGAGNHSYCRNPD 53
 | | | | | : | | | | | : | | | | |
 Db 103 CRYGNGKYYRGTVSKTKGLTCOKW-----SAETPHKPRFSPDENPSEGLDQNYCRNPD 156
 | | | | | : | | | | | : | | | | |
 Qy 54 EDPKGPWCYVSGEAGVPEKR--PCEDLRQPE 82
 | | | | | : | | | | | : | | | | |
 Db 157 NDPKGPWCYTM-----PEVRYEYCEIIQCED 183

Search completed: April 7, 2003, 09:16:56
 Job time : 13.0686 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:25 ; Search time 64.9023 Seconds

(without alignments)
768.284 Million cell updates/sec

Title: US-10-057-951-2_COPY_22_263

Sequence: 1 SGGCFWMDNGHLYREDQTSFA.....PVDQEGSTPLMGAGATPGCA 242

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacterioplasmid:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 1289 | 100.0 | 263 | 4 Q96FE7 | Q96FE7 homo sapien |
| 2 | 1286 | 99.8 | 263 | 4 O00318 | O00318 homo sapien |
| 3 | 167 | 13.0 | 562 | 6 O8S023 | O8S023 sus scrofa |
| 4 | 154.5 | 12.0 | 516 | 4 O9BU99 | O9BU99 homo sapien |
| 5 | 150 | 11.6 | 653 | 11 O8VC54 | O8VC54 mus musculu |
| 6 | 147.5 | 11.4 | 559 | 11 O91VP2 | O91VP2 mus musculu |
| 7 | 142 | 11.0 | 810 | 4 Q15146 | Q15146 homo sapien |
| 8 | 136.5 | 10.6 | 420 | 13 O90504 | O90504 epistretus |
| 9 | 135 | 10.5 | 385 | 5 Q25101 | Q25101 hermannia m |
| 10 | 135 | 10.5 | 812 | 11 O9R0W3 | O9R0W3 ratuus norv |
| 11 | 134.5 | 10.4 | 704 | 13 O90665 | O90665 gallus gall |
| 12 | 133.5 | 10.4 | 716 | 11 P70521 | P70521 ratuus norv |
| 13 | 131.5 | 10.2 | 313 | 13 O9PU78 | O9PU78 crocodylus |
| 14 | 131 | 10.2 | 154 | 4 O96588 | O96588 homo sapien |
| 15 | 131 | 10.2 | 608 | 13 O9PTW7 | O9PTW7 struthio ca |
| 16 | 130.5 | 10.1 | 716 | 11 Q91XG8 | Q91XG8 mus musculu |

| | | | | | |
|----|-------|------|------|-----------|---------------------|
| 17 | 130.5 | 10.1 | 717 | 13 P70006 | P70006 xenopus lae |
| 18 | 129 | 10.0 | 616 | 6 O97507 | O97507 sus scrofa |
| 19 | 128.5 | 10.0 | 806 | 6 O18783 | O18783 macropus eu |
| 20 | 127 | 9.9 | 157 | 6 O9TV48 | O9TV48 bos taurus |
| 21 | 127 | 9.9 | 395 | 4 O9BZW1 | O9BZW1 homo sapien |
| 22 | 126 | 9.8 | 716 | 13 Q91691 | Q91691 xenopus lae |
| 23 | 123 | 9.5 | 728 | 6 O9BH09 | O9BH09 felle silve |
| 24 | 119.5 | 9.3 | 399 | 4 O96GL8 | O96GL8 homo sapien |
| 25 | 119.5 | 9.3 | 420 | 4 O9BTP9 | O9BTP9 homo sapien |
| 26 | 119.5 | 9.3 | 560 | 4 O14520 | O14520 homo sapien |
| 27 | 119 | 9.2 | 334 | 6 O46507 | O46507 papio hamad |
| 28 | 118 | 9.2 | 812 | 13 Q91402 | Q91402 xenopus . he |
| 29 | 117.5 | 9.1 | 710 | 13 Q91402 | Q91402 xenopus . he |
| 30 | 114.5 | 8.9 | 113 | 4 O9UR75 | O9UR75 mus musculu |
| 31 | 114 | 8.8 | 113 | 4 O9UR75 | O9UR75 mus musculu |
| 32 | 114 | 8.8 | 202 | 13 O90675 | O90675 gallus gall |
| 33 | 114 | 8.8 | 607 | 13 Q91001 | Q91001 gallus gall |
| 34 | 113.5 | 8.8 | 452 | 13 O90Y90 | O90Y90 gallus gall |
| 35 | 112.5 | 8.7 | 359 | 6 O8MWR1 | O8MWR1 canis fami |
| 36 | 112 | 8.7 | 567 | 4 O13208 | O13208 homo sapien |
| 37 | 112 | 8.7 | 594 | 5 P91823 | P91823 caenorhadi |
| 38 | 112 | 8.7 | 1145 | 5 O9BKL8 | O9BKL8 drosophila |
| 39 | 111.5 | 8.7 | 397 | 11 O35727 | O35727 mus musculu |
| 40 | 111 | 8.6 | 726 | 13 Q90978 | Q90978 gallus gall |
| 41 | 110.5 | 8.6 | 685 | 5 Q24488 | Q24488 drosophila |
| 42 | 110 | 8.5 | 2358 | 16 O91LV8 | O91LV8 streptomyce |
| 43 | 108 | 8.4 | 215 | 13 O42341 | O42341 gallus gall |
| 44 | 106 | 8.2 | 378 | 13 O90WP0 | O90WP0 tracheomyx |
| 45 | 105.5 | 8.2 | 709 | 13 Q90ZNE | Q90ZNE brachydanio |

ALIGNMENTS

RESULT 1

Q96FE7 PRELIMINARY; PRT; 263 AA.
ID Q96FE7
AC Q96FE7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Unknown (protein for MGC:17330).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011049; AB011049.1; -.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle_1.
DR ProDom; PD000395; Kringle_1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR PROSITE; PS50070; KRINGLE_2; 1.
SQ
SEQUENCE 263 AA; 28234 MW; 197C3EE888FA242 CRC64;

Query Match 100.0%; Score 1289; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 9.5e-109;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFWMDNGHLYREDQTSFA...PVDQEGSTPLMGAGATPGCA 60
DB 22 SGGCFWMDNGHLYREDQTSFA...PVDQEGSTPLMGAGATPGCA 81
QY 61 CYVSGEAGVEKPCEDLRCPTTSQALPAFTTEIOEASEGCADEVQVAFANALPARS 120
DB 82 CYVSGEAGVEKPCEDLRCPTTSQALPAFTTEIOEASEGCADEVQVAFANALPARS 141
QY 121 EAAAVQPVGISQRYVMSKREKDKLGTGLVIGITMVIIAIGAGIIVGYSYKRGKDLK 180

Db 142 EAAAVPVIGISORVRNNSKEKDLGLGVIGITMMVILIAAGIILGISYKRGDK 201
QY 181 EOHDOKVCEREMORTITPLSAFTNPCTCEIVDEKTVVHTSQTPVDPOEGSTPLMGQAATP 240
Db 202 EOHDOKVCEREMORTITPLSAFTNPCTCEIVDEKTVVHTSQTPVDPOEGSTPLMGQAATP 261
QY 241 GA 242
Db 262 GA 263

RESULT 2
ID 000318 PRELIMINARY; PRT; 263 AA.
AC 000318;
DT 01-JUN-1997 (TREMBlrel. 04, Created)
DT 01-JUN-1997 (TREMBlrel. 04, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE WUSGC:DJ515N1.2 protein.
GN WUSGC:DJ515N1.2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_Taxid=9606;
OK NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Du Z., Scheet P., Harper M.;
RT "The sequence of H. sapiens PAC clone RP3-515N1."
RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AC002073; AAB54054.1; -.
DR HSSP: P00749; 1KDU.
DR InterPro: IPR000001; Kringle.
DR Pfam: PF00051; Kringle; 1.
DR PRINTS: PR00018; KRINGLE; FALSE_NEG.
DR PRODOM: PD000395; Kringle; 1.
DR SMART: SM00130; KR; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
SQ SEQUENCE 263 AA; 28248 MM; 197C3EEBE854A242 CRC64;

Query Match 99.8%; Score 1286; DB 4; Length 263;
Best Local Similarity 99.6%; Pred. No. 1.8e-108;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFMDNGHLVREDQTSAPGLRCINWLDASGLASAPVSGAGNHSYCRNDEDEPRGPW 60
Db 22 SGGCFMDNGHLVREDQTSAPGLRCINWLDASGLASAPVSGAGNHSYCRNDEDEPRGPW 81
QY 61 CYSVSGAGVPEKRPCEDLRCPEPTTSQALPAFTTEIOEASEGPGADEVQYFAPANALPARS 120
Db 82 CYSVSGAGVPEKRPCEDLRCPEPTTSQALPAFTTEIOEASEGPGADEVQYFAPANALPARS 141
QY 121 EAAAVPVIGISORVRNNSKEKDLGLGVIGITMMVILIAAGIILGISYKRGDK 180
Db 142 EAAAVPVIGISORVRNNSKEKDLGLGVIGITMMVILIAAGIILGISYKRGDK 201
QY 181 EOHDOKVCEREMORTITPLSAFTNPCTCEIVDEKTVVHTSQTPVDPOEGSTPLMGQAATP 240
Db 202 EOHDOKVCEREMORTITPLSAFTNPCTCEIVDEKTVVHTSQTPVDPOEGSTPLMGQAATP 261
QY 241 GA 242
Db 262 GA 263

RESULT 3
ID 085023 PRELIMINARY; PRT; 562 AA.
AC 085023;

DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
NCBI_Taxid=9823;
OK NCBI_Taxid=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ENAMEL ORGAN;
RA Ding Y., Xue J., Bartlett J.D.;
RT "T-plasminogen activator in tooth tissues."
RL Submitted (Mar-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF364605; AA00297.1; -.
SQ SEQUENCE 562 AA; 63668 MM; P9B6B4C77CB101E8 CRC64;

Query Match 13.0%; Score 167; DB 6; Length 562;
Best Local Similarity 36.6%; Pred. No. 1e-06;
Matches 41; Conservative 12; Mismatches 33; Indels 26; Gaps 6;

QY 4 CFMDNGHLVREDQTSAPGLRCINWLDASGLASAPVSGAGNHSYCRNDEDEPRGPW 54
Db 127 CYBDQGITRYGTSTSTESGAEQVNW--NTSGLASPPYNGRRPDAYKLGHNHNYCRNPDK 184
QY 55 DPRGPGCVV-SGAGVPEKRPCEDLRCPEPTTSQALPAFTTEIOEASEGPGAD 105
Db 185 DSK-PWCYIFPAKATPSD-----FC-----STPACTKEKEBCTGTGGLD 222

RESULT 4
ID 09B099 PRELIMINARY; PRT; 516 AA.
AC 09B099;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Similar to plasminogen activator, tissue.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_Taxid=9606;
OK NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.
DR EMBL: BC002795; AA002795.1; -.
DR HSSP: P00750; 1A5H.
DR MEROPS: S01.232; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF00051; Kringle; 2.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR PRODOM: PD000395; Kringle; 2.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00001; EGF-like; 1.
DR SMART: SM00130; KR; 2.
DR SMART: SM00020; TRY_PSP; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01166; EGF_2; 1.
DR PROSITE: PS00021; KRINGLE_1; UNKNOWN_2.
DR PROSITE: PS00070; KRINGLE_2; 2.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
DR PROSITE: PS0134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.

KW EGF-like domain; Glycoprotein; Hydrolyase; Serine protease.
SQ SEQUENCE 516 AA; 57370 MW; BAB1901FDC36800 CRC64;

Query Match 12.0%; Score 154.5; DB 4; Length 516;
Best Local Similarity 39.6%; Pred. No. 1.2e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFWDNGHLVREDOTSPAPGLRCLNWLDAQSGLASAPV-----GAGNHSYCRNPDE 54
DB 81 CREDQISYRGTWSTASGABCTNM--NSALAAQPYSGRPDALRGLGNHNYCRNPDR 138
QY 55 DRGPMCVYSGAGVPEKRPCEDLRCPEFTTS 85
DB 139 DSK-FWCTVF--KAGKYSSEFCSTPACSEGN 167

RESULT 5
Q8VCS4 PRELIMINARY; PRT; 653 AA.

AC Q8VCS4;
DT 01-MAR-2002 (TREMblrel, 20, Created)
DT 01-MAR-2002 (TREMblrel, 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel, 21, Last annotation update)
DE Hypothetical 70.6 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAUSBERG R.;
RA Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC019376; AAH19376.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001438; EGF_11.
DR InterPro; IPR000083; Fibronectin.
DR InterPro; IPR000562; FN_Type_11.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle_1.
DR Pfam; PF00089; trypsin_1.
DR PRINTS; PR00722; CHYMOTRPSIN.
DR PRINTS; PR0010; EGFBCLOD.
DR PRINTS; PR0013; FNTYPE11.
DR PRINTS; PR0018; TRYPE11.
DR ProDom; PD000395; Kringle_1.
DR ProDom; PD000995; FN_Type_11; 1.
DR SMART; SM00181; EGF_2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYF_SPC; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS01253; FIBRONECTIN_1; UNKNOWN_1.
DR PROSITE; PS00023; FIBRONECTIN_2; UNKNOWN_1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02440; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hypothetical proteain.
SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;

Query Match 11.6%; Score 150; DB 11; Length 653;
Best Local Similarity 33.8%; Pred. No. 4.3e-05;
Matches 47; Conservative 18; Mismatches 44; Indels 32; Gaps 8;

QY 4 CFWDNGHLVREDOTSPAPGLRCLNWLDAQSGLASAPVGAAGNHSYCRNPDE 55
DB 283 CFLANGTEYGVAVSTASGLSCIANNSDLYVOELHVDV--VAAVLLGLDPHAYCRNPDK 341
QY 56 PRGPMCVYSGAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASGPGADEVQVAPANA 115
DB 342 ER-FWCTVVDNMLSWE-----YCRITACESTL-----ARVHSGSP-----EILA---A 380
QY 116 LPARSEAAVQVPIYIGSQVRV 136
DB 381 LP--ESAPAVRPTCGRHKKR 399

RESULT 6
Q91VP2 PRELIMINARY; PRT; 559 AA.

AC Q91VP2;
DT 01-DEC-2001 (TREMblrel, 19, Created)
DT 01-DEC-2001 (TREMblrel, 19, Last sequence update)
DT 01-JUN-2002 (TREMblrel, 21, Last annotation update)
DE Similar to plasminogen activator, tissue.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRST TUMOR;
RA STRAUSBERG R.;
RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011256; AAH11256.1; -.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000083; Fibronectin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle_2.
DR Pfam; PF00089; trypsin; 1.
DR ProDom; PD000395; Kringle_2.
DR SMART; SM00181; EGF_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS01253; FIBRONECTIN_1; UNKNOWN_1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS02440; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolyase; Serine protease.
SQ SEQUENCE 559 AA; 63122 MW; 8CCEB2BDB94514D9 CRC64;

Query Match 11.4%; Score 147.5; DB 11; Length 559;
Best Local Similarity 37.0%; Pred. No. 5.9e-05;
Matches 34; Conservative 11; Mismatches 34; Indels 13; Gaps 4;

QY 4 CFWDNGHLVREDOTSPAPGLRCLNWLDAQSGLASAPV-----GAGNHSYCRNPDE 54
DB 124 CFEEGGLTRKGTWSTASGABCTNM--NSVYLAKPTNARPNATKLGAGNHYCRNPDR 181
QY 55 DRGPMCVYSGAGVPEKRPCEDLRCPEFTTSQ 86
DB 182 DLK-FWCTVF--KAGKYTEFCSTPACPKGKE 211

RESULT 7
Q15146 PRELIMINARY; PRT; 810 AA.

AC Q15146;
DT 01-NOV-1996 (TREMblrel, 01, Created)
DT 01-NOV-1996 (TREMblrel, 01, Last sequence update)
DT 01-JUN-2002 (TREMblrel, 21, Last annotation update)

```

DE Plasminogen precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Bryone M.U., Chapman C.G., Dodd I., Carey J.E., Lawrence G.M.P.,
RA Mitchell D., Robinson J.H.;
RT "Expression of recombinant human plasminogen and aglycoplasminogen in
RT HeLa cells.";
RL Fibrinolysis 0:0-0(1991).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: M74220; AAA36451.1; -.
DR HSP: P00747; 2PK4.
DR MEROPS: S01.233; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00051; Kringle_5.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle_5.
DR SMART: SM00130; KR; 5.
DR SMART: SM00473; PAN_Ap; 1.
DR PROSITE: PS00020; TRYP_SPC; 1.
DR PROSITE: PS00021; KRINGLE_1; 5.
DR PROSITE: PS50070; KRINGLE_2; 5.
DR PROSITE: PS50240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; 1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
DR KX Hydroxylase; Serine protease; Signal.
FT SIGNAL 1
FT CHAIN 1
SQ SEQUENCE 810 AA; 90555 MW; B05C7D4BD0D020B3C CRC64;

Query Match 11.0%; Score 142; DB 4; Length 810;
Best Local Similarity 37.2%; Pred. No. 0.0003;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 4 CFWDNGHLYREDQTSAPAGRLCLNWLDA-----QSGLASAPVSGAGNHSYCRNPDDPRG 58
DB 103 CKTGNGKNYRGTMSTKRGITCQMSSTSPHPRFSPATHPSEGL-BENYCRNPDNDPOG 161
QY 59 PWCYVSGEAGVPEKR--PCEDLRCP 82
DB 162 PWCYTTD---PEKKYDYCDILECEE 183

```

RESULT 8

```

ID Q90504 PRELIMINARY; PRT; 420 AA.
AC Q90504;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Thrombin.
OS Eptatretus scoulli (Pacific hagfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;
OC Myxiniidae; Eptatretinae; Eptatretus.
OX NCBI_TaxId=7765;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Bantfield D.K., Macgillivray R.T.;
RA Bantfield D.K., Macgillivray R.T.;
RT "Partial characterization of vertebrate prothrombin cDNAs

```

```

RT amplification and sequence analysis of the B chain of thrombin from
RT nine different species.";
RT Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=9423694; PubMed=7513365;
RA Bantfield D.K., Irwin D.M., Walz D.A., Macgillivray R.T.;
RT "Evolution of prothrombin: isolation and characterization of the cDNAs
RT encoding chicken and hagfish prothrombin.";
RT J. Mol. Evol. 38:177-187(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Bantfield D.K.;
RT Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: M81393; AAA21620.1; -.
DR HSP: P00734; 1UV6.
DR MEROPS: S01.217; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00051; Kringle_1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle_1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; TRYP_SPC; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS50070; KRINGLE_2; 1.
DR PROSITE: PS50240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
DR KX Hydroxylase; Serine protease.
SQ SEQUENCE 420 AA; 47888 MW; 64522AA21A57B67A CRC64;

```

```

Query Match 10.6%; Score 136.5; DB 13; Length 420;
Best Local Similarity 26.5%; Pred. No. 0.00041;
Matches 39; Conservative 16; Mismatches 77; Indels 15; Gaps 4;

```

```

QY 4 CFWDNGHLYREDQTSAPAGRLCLNWLDAQSGLASAPVSGAG-NHSYCRNPDDPRGPNCY 62
DB 17 CYRREGDYRGDNLITWTGKPLPWRGSYSNLPDQFTTAGLTSNYCRNPDDGDSRGWCY 76
QY 63 VSGEAGVP---EKPCEDLRCPETTSQALPAFTTEIQEASBSPGADDEVAPANALPA 118
DB 77 TYGVEGTVDVDCOLNACE-----SGDIFVGTDEVQLSGRSEGAAEKTLFFNPKTFGN 129
QY 119 RSEAAAVPVIGISORVRNMSKEDL 145
DB 130 GEEGCKRPWFELQK---NDRSEDEL 153

```

RESULT 9

```

ID Q25101 PRELIMINARY; PRT; 385 AA.
AC Q25101;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Serine proteinase.
OS Herdmania momus.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Fyuridae; Herdmania.
OX NCBI_TaxId=7733;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CURVATA;
RA Arnold J.M., Kennett C., Lavin M.F.;

```


RT "Transient expression of a novel serine protease in the ectoderm of
 RL the ascidian *Herdmania momus* during development.";
 CC Dev. Genes Evol. 206:455-463(1997).
 -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL: U63517; AAB66650.1; -.
 DR HSSP: P00763; IDPO.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Ser. protease_Try.
 DR Pfam: PF00089; kringle; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR PRODOM: PD000395; Kringle; 1.
 DR SMART: SM00130; KR; 1.
 DR SMART: SM00020; TRYPSIN; 1.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS00070; KRINGLE_2; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR KEGG: K00001; Serine protease.
 SQ SEQUENCE 385 AA; 42935 MW; BFBID05D5232E6A0 CRC64;

Query Match 10.5%; Score 135; DB 5; Length 385;
 Best Local Similarity 20.2%; Pred. No. 0.0005;
 Matches 65; Conservative 44; Mismatches 121; Indels 92; Gaps 12;

QY 4 CF-WDNGHLYREDQTPAPGRLCNLWLDAS-----GLASAVSGAGNHSYCRNDDDEP 56
 DB 23 CFIENNESISQGAISRLTGETEQSW-DLQTPHKHKTYSNYPNSGLAGNNYCRNDQDM 81
 QY 57 RGFVCYVSGE-----AGVPEKPCEDLRCPETT---SQALPAFTTEIOEASG 101
 DB 82 RGFVCYVSGE-----AGVPEKPCEDLRCPETT---SQALPAFTTEIOEASG 101
 QY 102 P-----GADEVQ-----VFAPNALPARSEAAAVQPV 128
 DB 142 PLHIVGGTTVTHGSIPOVSLRLKRELHFCGSIINRWILFAHCHIRPQPKYLA 201
 QY 129 IGISORVNMSEKEDLGLGYVL-----GITMNVIIAIGAGIILGYSK 174
 DB 202 LGYVDALQVDFSEMK-----VGFRLFNHKKYNPATENDITLTKMDTSSISATIFQGSVP 257
 QY 175 RGDLDKQHDQVC-----EREMORITLPLSAFTNPTCEIVDEKTV---VVH 218
 DB 258 PPAKVPAAKSKIIIVSGWEDTKGTQVKNQVTLFVMSF--KLCKKLYSKYVGAAPVFK 315
 QY 219 TSGTTPVDPQGSTPLMGQACTP 240
 DB 316 TSLCAAYKKGKDSQGGDSGGP 337

RESULT 10
 Q980W3 PRELIMINARY; PRT; 812 AA.
 AC Q980W3;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Plasmidogen protein precursor (EC 3.4.21.7).
 GN PLASMINOGEN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Bangert K., Johnsen A.H., Thorsen S.;
 RT "Rat plasmidogen: cDNA and gene structure."
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=91250378; Pubmed=1645711;
 RA Kanalaas J.J., Makker S.P.;
 RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
 RL receptor site for plasmidogen.";
 RT J. Biol. Chem. 266:10825-10828 (1991).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL: AJ242649; CAB46014.1; -.
 DR HSSP: P00747; 1PMK.
 DR MEROPS: S01.233; -.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR003014; PAN.
 DR InterPro: IPR003609; Pan_app.
 DR InterPro: IPR001254; Ser. protease_Try.
 DR InterPro: IPR001400; Somatostatin.
 DR Pfam: PF00051; kringle; 5.
 DR Pfam: PF00024; PAN; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR PRODOM: PD000395; Kringle; 5.
 DR SMART: SM00130; KR; 4.
 DR SMART: SM00473; PAN AP; 1.
 DR SMART: SM00020; TRYPSIN; 1.
 DR PROSITE: PS00021; KRINGLE_1; 5.
 DR PROSITE: PS00070; KRINGLE_2; 5.
 DR PROSITE: PS50240; SOMATOSTATIN_2; UNKNOWN_1.
 DR PROSITE: PS00134; TRYPSIN_DOM; 1.
 DR PROSITE: PS00135; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR KEGG: K00001; Serine protease; Signal.
 KM Hydrolyase; Serine protease; Signal.
 FT SIGNAL
 FT CHAIN
 SQ SEQUENCE 812 AA; 90535 MW; 8C703C51410EBC9E CRC64;

Query Match 10.5%; Score 135; DB 11; Length 812;
 Best Local Similarity 30.9%; Pred. No. 0.0013;
 Matches 34; Conservative 14; Mismatches 48; Indels 14; Gaps 5;
 QY 4 CFWDNGHLYREDQTPAPGRLCNLWLDAS-----QSLASAPVSGAGNHSYCRNDDDEP 58
 DB 376 CYQNGKSYKRTSTTMTGKKCSWVSMTPHSKSTPAPNPDHGL-EMNYCRNDDQDG 434
 QY 59 PWCYSGEAGVPEKR--PCEDLRCPETTSQLPAPFTTEIOEASGPGADE 106
 DB 435 PWCFTTD-----PSVMEYCNLKRCSFTGGV--AESALVPQVSAPOTSE 478

RESULT 11
 Q90865 PRELIMINARY; PRT; 704 AA.
 AC Q90865;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hepatocyte growth factor-like/macrophage stimulating protein.
 GN HGF1/MSP.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=96029010; Pubmed=7554499;
 RA Thery C., Sharpe M.J., Batley S.J., Stern C.D., Gherardi E.;
 RT "Expression of HGF/SF, HGF1/MSP and c-met suggests new functions
 during early chick development.";

```

RL   Dev. Genet. 17:90-101(1995).
CC   -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE
CC     TRYPSIN FAMILY.
DR   EMBL: X84043; CA5862.1; -.
DR   HSSP: P00747; ICEA.
DR   MEROPS: S01.977; -.
DR   InterPro: IPR001314; Chymotrypsin.
DR   InterPro: IPR000001; Kringle.
DR   InterPro: IPR003014; PAN.
DR   InterPro: IPR003609; Pan_app.
DR   InterPro: IPR001254; Ser.protease_Try.
DR   Pfam: PF00051; kringle; 4.
DR   Pfam: PF00024; PAN; 1.
DR   Pfam: PF00089; trypsin; 1.
DR   PRINTS: PR00722; CHYMOTRYPSIN.
DR   PRINTS: PR00018; KRINGLE.
DR   ProDom: PD000395; Kringle; 4.
DR   SMART: SM00130; KR; 4.
DR   SMART: SM00473; PAN_Ap; 1.
DR   SMART: SM00020; TRYF_Spc; 1.
DR   PROSITE: PS00021; KRINGLE_1; 4.
DR   PROSITE: PS50070; KRINGLE_2; 4.
DR   PROSITE: PS50240; TRYPSIN_DOM; 1.
KM   Hydrolyase; Serine protease.
SQ   SEQUENCE 704 AA; 79341 MW; CAB0D8CC41367C37 CRC64;

Query Match          10.4%; Score 134.5; DB 13; Length 704;
Best Local Similarity 28.8%; Pred. No. 0.0012;
Matches 32; Conservative 11; Mismatches 45; Indels 23; Gaps 4;

OY 4 CFWDNGLYREDQTSPPAPGLRCLNW--LDAQSLGASAPVSGAGNHSYCRNPDED 55
DB 108 CIVANGSYRGTRDTERGLRCQHW-----QATPHHRFLPSIRKLNENYCRNPDRD 161
OY 56 PRGFWCYV-----SGEAVPEKRPCEDLRCPETTSQALPAFTTEIOEASE 100
DB 162 KRGPWCYIVDPVVRHQSGLI---KKCEDAVCMTCNGEDYRGVHDHTESGTE 209

RESULT 12
P70521 PRELIMINARY; PRT; 716 AA.
AC P70521;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Macrophage stimulating protein precursor.
OS MSP.
OC Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=97011126; PubMed=8658136;
RA Takasu N., Suda T.;
RA Ohshtiro K., Yamada A., Matsuno K., Ezaki T., Sakamoto O., Hamaguchi I.,
RT "Molecular cloning of Rat Macrophage-stimulating protein and its
RT involvement in the Male Reproductive System.";
RT Biochem. Biophys. Res. Commun. 227:273-280(1996).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE
CC     TRYPSIN FAMILY.
DR   EMBL: X95096; CA64473.1; -.
DR   HSSP: P00747; IREN.
DR   MEROPS: S01.975; -.
DR   InterPro: IPR001314; Chymotrypsin.
DR   InterPro: IPR000001; Kringle.
DR   InterPro: IPR003014; PAN.
DR   InterPro: IPR003609; Pan_app.
DR   InterPro: IPR001254; Ser.protease_Try.
DR   Pfam: PF00051; kringle; 4.
DR   Pfam: PF00024; PAN; 1.

```

```

DR   Pfam: PF00089; trypsin; 1.
DR   PRINTS: PR00722; CHYMOTRYPSIN.
DR   PRINTS: PR00018; KRINGLE.
DR   ProDom: PD000395; Kringle; 4.
DR   SMART: SM00130; KR; 4.
DR   SMART: SM00473; PAN_Ap; 1.
DR   SMART: SM00020; TRYF_Spc; 1.
DR   PROSITE: PS00021; KRINGLE_1; 4.
DR   PROSITE: PS50070; KRINGLE_2; 4.
DR   PROSITE: PS50240; TRYPSIN_DOM; 1.
KM   Hydrolyase; Serine protease; Signal.
FT SIGNAL 1
FT SIGNAL 31
SQ SEQUENCE 716 AA; 80733 MW; 06B7DF3EF56D921F CRC64;

Query Match          10.4%; Score 133.5; DB 11; Length 716;
Best Local Similarity 29.1%; Pred. No. 0.0015;
Matches 41; Conservative 9; Mismatches 48; Indels 43; Gaps 6;

OY 4 CFWDNGLYREDQTSPPAPGLRCLNW--LDAQSLGASAPVSGAGNHSYCRNPDEPRGPW 60
DB 110 C1MDNGASRGTVARTADQLPQAWSRFRPNDHKYTPPKNGL-EENFCRNPDPGPRGPW 168
OY 61 CYVS-----GEGVPEK-RPCE--DLRCPET----- 83
DB 169 CYTNRSVRFQSGIKSCREAVCWNGEDYRGVEVDYTESGECORWDLQHPHSHFPAPE 228
OY 84 --TSQALPAFTTEIOEASEGP 102
DB 229 KPPDKALKNYCRNPDAISRP 249

RESULT 13
Q9PU78 PRELIMINARY; PRT; 313 AA.
AC Q9PU78;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hepatocyte growth factor-like protein (Fragment).
OS Crocodylus niloticus (Nile crocodile) (African crocodile).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodyliidae; Crocodylinae; Crocodylus.
OC NCBI_TaxID=8501;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=20022983; PubMed=10555283;
RA Hughes S., Zelus D., Mouchiroud D.;
RT "Warm-blooded Isochoire structure in nile crocodile and turtle.";
RL Mol. Biol. Evol. 16:1521-1527(1999).
DR   EMBL: AJ011396; CAB56422.1; -.
DR   HSSP: P00747; IHPJ.
DR   MEROPS: S01.977; -.
DR   InterPro: IPR000001; Kringle.
DR   InterPro: IPR001254; Ser.protease_Try.
DR   Pfam: PF00051; kringle; 2.
DR   Pfam: PF00089; trypsin; 1.
DR   PRINTS: PR00018; KRINGLE.
DR   ProDom: PD000395; Kringle; 1.
DR   SMART: SM00130; KR; 1.
DR   SMART: SM00020; TRYF_Spc; 1.
DR   PROSITE: PS00021; KRINGLE_1; 1.
DR   PROSITE: PS50070; KRINGLE_2; 1.
DR   PROSITE: PS50240; TRYPSIN_DOM; 1.
KM   Hydrolyase; Serine protease.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 313
SQ SEQUENCE 313 AA; 34793 MW; 8E084704958B5AA2 CRC64;

Query Match          10.2%; Score 131.5; DB 13; Length 313;
Best Local Similarity 29.8%; Pred. No. 0.00079;
Matches 28; Conservative 11; Mismatches 42; Indels 13; Gaps 4;

```

QY 4 CFWDNGHLYREDQTSAPGLRCLNWLDAQSLAS-APVSGAGNH--SYCNPDEDEPRGP 59
DB 30 CYHNGELHYRGHTSKTRKGVTCQKWSQSPHVPQISPTTHPAALHDEYCNPNDSHGP 89
QY 60 WCYVSGE-----AGVPEKRPCEDLRCPEETSQA 87
DB 90 WCYTDPRTPFDYCGI---KPCAGDKIPSVLENA 120
RESULT 14
096SE8 PRELIMINARY; PRT; 154 AA.
ID 096SE8
AC 096SE8; 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Urokinase-type plasminogen activator amino-terminal fragment.
GN ATP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OK NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Fu J., Bai X., Ruan C.;
RT "Cloning and expression of the amino-terminal fragment of human urokinase-type plasminogen activator";
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Bai X., Fu J., Wang W., Xi X., Ruan C.;
RT "Overexpression of the amino-terminal fragment of human urokinase-type plasminogen activator in breast cancer cells results in decreased tumor invasion, growth and angiogenesis";
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY029537; AAK38734.1; -
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00051; kringle_1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Kinase.
SQ SEQUENCE 154 AA; 17305 MW; A3CCP2FCFFS0S572 CRC64;
Query Match 10.2%; Score 131; DB 4; Length 154;
Best Local Similarity 40.3%; Pred. No. 0.00036;
Matches 27; Conservative 7; Mismatches 25; Indels 8; Gaps 2;
QY 4 CFWDNGHLYREDQTSAPGLRCLNWLDAQSLASAPVSGAGNH--SYCNPDEDEPRGP 56
DB 70 CLEGNHFRFGKASTDTMRPCLPMNSATVTLQOTTHAFHRSALQJLGNHNYCNPNP-DNR 128
QY 57 RGPWCYV 63
DB 129 RRPWCYV 135
RESULT 15
09PTW7 PRELIMINARY; PRT; 608 AA.
ID 09PTW7
AC 09PTW7; 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Prothrombin.
GN OSPT.
OS Struthio camelus (Ostrich).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Palaeognathae; Struthioniformes; Struthionidae;

OC Struthio.
OK NCBI_TaxID=8801;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=20579470; PubMed=11137455;
RA Frost C., Nade R., Oelofsen W., Muramoto K., Naganuma T., Ogawa T.;
RT "Purification and characterization of ostrich prothrombin";
RT Int. J. Biochem. Cell Biol. 32:1151-1159(2000).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.
DR EMBL; AB028871; BA89046.1; -
DR HSSP; P00734; IUVS.
DR MEROPS; S01.217; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR002383; GLA blood.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser protease_Try.
DR InterPro; IPR000284; Vitr_dep_GLA.
DR Pfam; PF00594; gla; 1.
DR Pfam; PF00051; Kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00001; GLABLOOD.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00069; GLA; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Ser; 1.
DR PROSITE; PS00011; GH CARBOXYLATION; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolyase; serine protease.
SQ SEQUENCE 608 AA; 69392 MW; 11B974B9ABE54EA2 CRC64;
Query Match 10.2%; Score 131; DB 13; Length 608;
Best Local Similarity 31.3%; Pred. No. 0.002;
Matches 41; Conservative 12; Mismatches 68; Indels 10; Gaps 5;
QY 2 GGCWDNGHLYREDQTSAPGLRCLNWLDAQSLASAPVSGAGN--HSYCNPEDEPRGP 59
DB 107 GNCSTDLGQVYRGTLISHKSGTEQMTSKYRPHIPKNTTHHPNLINENYCNPNPNDEGP 166
QY 60 WCYVSGEAGVPEKRP---C-EDLRCPEETSQALPAFTTEIOEASEG--PGADEVGVFAP 112
DB 167 WCYTRDPTVPRRESCFIPVCGERTTVEFTFRVXPPASTPECGEKMLVAGTSTVTI-SG 225
QY 113 ANLPAASENA 123
DB 226 AKCLPMNSEEA 236

Search completed: April 7, 2003, 09:21:35
Job time : 66.9023 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:24 ; Search time 45.2807 Seconds
(without alignments)
712.151 Million cell updates/sec

Title: US-10-057-951-2_COPY_22_263
Perfect score: 1289
Sequence: 1 SGGCFWMDNGHLXREDQTPA.....PVDPEGSTRPLMGAGTPGA 242

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1 | 1289 | 100.0 | 263 | 20 | AAV05219 |
| 2 | 1289 | 100.0 | 263 | 20 | AAV05219 |
| 3 | 1289 | 100.0 | 263 | 20 | AAV05219 |
| 4 | 1286 | 99.8 | 263 | 23 | AAU06149 |
| 5 | 1280 | 99.3 | 263 | 21 | AAU06149 |
| 6 | 1275 | 98.9 | 263 | 22 | AAU06149 |
| 7 | 1063.5 | 82.5 | 286 | 20 | AAV05220 |
| 8 | 322 | 25.0 | 66 | 22 | ABR37905 |
| 9 | 322 | 25.0 | 66 | 22 | ABR37905 |
| 10 | 322 | 25.0 | 66 | 22 | ABR37905 |

| | | | | | | |
|----|-------|------|-----|----|----------|--------------------|
| 11 | 322 | 25.0 | 66 | 22 | AAV71037 | Human bone marrow |
| 12 | 322 | 25.0 | 66 | 22 | AAV18800 | Peptide #5234 enco |
| 13 | 322 | 25.0 | 66 | 22 | AAV13114 | Peptide #5351 enco |
| 14 | 322 | 25.0 | 66 | 23 | ABG40828 | Human peptide enco |
| 15 | 202 | 15.7 | 56 | 20 | AAV12397 | Human 5' EST score |
| 16 | 201 | 15.6 | 55 | 20 | AAV12397 | Human 5' EST score |
| 17 | 196 | 15.2 | 39 | 19 | AAV72641 | Nervous glia cell |
| 18 | 192 | 14.9 | 39 | 19 | AAV72640 | Nervous glia cell |
| 19 | 160.5 | 12.5 | 527 | 13 | AAV20217 | t-PA analogue expr |
| 20 | 160.5 | 12.5 | 527 | 13 | AAV20220 | t-PA analogue expr |
| 21 | 160.5 | 12.5 | 527 | 13 | AAV20218 | t-PA analogue expr |
| 22 | 160.5 | 12.5 | 527 | 13 | AAV20219 | t-PA analogue expr |
| 23 | 159.5 | 12.4 | 527 | 13 | AAV20223 | t-PA analogue expr |
| 24 | 159.5 | 12.4 | 527 | 13 | AAV20222 | t-PA analogue expr |
| 25 | 159.5 | 12.4 | 527 | 13 | AAV20221 | t-PA analogue expr |
| 26 | 157.5 | 12.2 | 527 | 19 | AAV54154 | t-PA mutant (N142S |
| 27 | 157.5 | 12.2 | 527 | 19 | AAV54157 | t-PA mutant (N142S |
| 28 | 156.5 | 12.1 | 439 | 16 | AAV68851 | Delta 2-89 tissue |
| 29 | 156.5 | 12.1 | 483 | 16 | AAV70889 | Human tissue PA va |
| 30 | 156.5 | 12.1 | 483 | 16 | AAV70877 | Human tissue PA va |
| 31 | 156.5 | 12.1 | 483 | 16 | AAV70878 | Human tissue PA va |
| 32 | 156.5 | 12.1 | 483 | 16 | AAV70879 | Human tissue PA va |
| 33 | 156.5 | 12.1 | 483 | 16 | AAV70880 | Human tissue PA va |
| 34 | 156.5 | 12.1 | 483 | 16 | AAV70881 | Human tissue PA va |
| 35 | 156.5 | 12.1 | 483 | 16 | AAV70882 | Human tissue PA va |
| 36 | 156.5 | 12.1 | 483 | 16 | AAV70883 | Human tissue PA va |
| 37 | 156.5 | 12.1 | 483 | 16 | AAV70884 | Human tissue PA va |
| 38 | 156.5 | 12.1 | 483 | 16 | AAV70885 | Human tissue PA va |
| 39 | 156.5 | 12.1 | 483 | 16 | AAV70886 | Human tissue PA va |
| 40 | 156.5 | 12.1 | 483 | 16 | AAV70887 | Human tissue PA va |
| 41 | 156.5 | 12.1 | 483 | 16 | AAV70888 | Human tissue PA va |
| 42 | 156.5 | 12.1 | 483 | 16 | AAV70890 | Human tissue PA va |
| 43 | 156.5 | 12.1 | 527 | 13 | AAV21598 | tPA variant - T103 |
| 44 | 156.5 | 12.1 | 527 | 14 | AAV44812 | Human tPA variant |
| 45 | 156.5 | 12.1 | 527 | 14 | AAV44816 | Human tPA variant |

ALIGNMENTS

| | | |
|----------|--|----------------------------|
| RESULT 1 | AAV05219 | standard; Protein; 263 AA. |
| ID | AAV05219 | standard; Protein; 263 AA. |
| XX | AAV05219 | |
| AC | AAV05219 | |
| DT | 17-JUN-1999 | (first entry) |
| XX | | |
| DE | Kringel protein sequence. | |
| XX | | |
| KW | Kringel; cancer; inflammation; autoimmunity; rheumatoid arthritis; AIDS; | |
| KW | CNS inflammation; cerebral degeneration; Alzheimer's disease; ashma; | |
| KW | Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis; | |
| KW | neurological abnormality; ischemia reperfusion injury; ischaemic injury; | |
| KW | cardiovascular disease; kidney disease; liver disease; aplastic anaemia; | |
| KW | myocardial infarction; hypotension; hypertension; allergy; infection; | |
| KW | myelodysplastic syndrome; haematologic abnormality; diagnosis; therapy; | |
| KW | male pattern baldness. | |
| OS | Homo sapiens. | |
| XX | | |
| PN | WO911788-A1. | |
| XX | | |
| PD | 11-MAR-1999. | |
| XX | | |
| PF | 02-SEP-1998; | 98WO-US18270. |
| XX | | |
| PR | 01-SEP-1998; | 98US-0144889. |
| XX | | |
| PR | 02-SEP-1997; | 97US-0056032. |
| XX | | |
| PA | (SMK) SMITHKLINE BEECHAM CORP. | |
| XX | | |
| PI | Albone EF, Kikly KK; | |

XX WPI; 1999-214707/18.
DR N-PSDB; AAX28354.
DR

DR N-PSDB; AAX28354.

PT New kringiel polypeptides and polynucleotides

PS Claim 1; Page 31-32; 42pp; English.

CC This sequence is a Kringle polypeptide of the invention.

The kringle polypeptides (I) are used to screen for agonists and antagonists. Agonists are used to treat subjects in need of enhanced activity or expression of (I). Antagonists are used to treat subjects having need to inhibit the activity or expression of (I). The methods can be used to treat conditions such as cancer, inflammation, autoimmunity, allergy, asthma, rheumatoid arthritis, CNS inflammation, cerebellar degeneration, Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis, head injury damage and other neurological abnormalities, ischaemia reperfusion injury, cardiovascular disease, kidney disease, liver disease, ischaemic injury, myocardial infarction, hypotension, AIDS, myelodysplastic syndromes and other haematologic abnormalities, aplastic anaemia, male pattern baldness, and bacterial, fungal, protozoan and viral infections. The kringle polypeptides may also be used to generate antibodies. Determining the presence or absence of mutations in, and analysing for the presence or absence of expression of, kringle polynucleotides can be used to diagnose a disease or susceptibility to a disease related to expression or activity of kringle proteins. The polynucleotides may also be used for chromosome identification, and mapping.

Determining the presence or absence of mutations in, and analysing for the presence or absence of expression of, kringel polynucleotides can be used to diagnose a disease or susceptibility to a disease related to expression or activity of kringel proteins. The polynucleotides may also be used for chromosome identification, and mapping.

SQ Sequence 263 AA;

| | | | | |
|---------------------------|---------|---------------------|-----------|-------------|
| Query Match | 100.0% | Score 1289; | DB 20; | Length 263; |
| Best Local Similarity | 100.0%; | Pred. No. 4.6e-114; | | |
| Matches 242; Conservative | 0; | Mismatches 0; | Indels 0; | Gaps 0; |

| | | | | | |
|----|-----|---|--------------------------------|------------------------|-----|
| Qy | 1 | SGGCGMNDGHHYREDOQTS | PAGCLRCLMNLMDQSGLAS | PVSGAAGNHSYCRNPDEDPGPM | 60 |
| Db | 22 | SGGCGMNDGHHYREDOQTS | PAGELRCMLNLDQSGLAS | PVSGAAGNHSYCRNPDEDPGPM | 81 |
| Qy | 61 | CYVSGEAGVPEKRPCEDRCEPTTSQALPAFTTEIOEASBPQADEVOVFAPNALPARS | | | 120 |
| Db | 82 | CYVSGEAGVPEKRPCEDRCEPTTSQALPATTEIOEASBPQADEVOVFAPNALPARS | | | 144 |
| Qy | 121 | EAAGVOPVIGISQVRVMSKSKKDLGTLGYVLGITMVI | IIAIGAGIILGYSYRGDKL | | 180 |
| Db | 142 | EAAGVOPVIGISQVRVMSKSKKDLGTLGYVLGITMVI | IIAIGAGIILGYSYRGDKL | | 201 |
| Qy | 181 | EOHODKVCEREMQRTTLPLSAFTNPCTCI | VEBKTVVHTISQTPVDPQBSGFLPLMGQGT | | 240 |
| Db | 202 | EOHODKVCEREMQRTTLPLSAFTNPCTCI | VEBKTVVHTISQTPVDPQBSGFLPLMGQGT | | 261 |

| | |
|----------|-------------------------------------|
| RESULT 2 | |
| AAW87769 | |
| ID | AAW87769 standard; Protein; 263 AA. |

AC AAW87769;

DT 29-MAR-1999 (first entry)

Human tissue plasminogen activator-like protease t-PA/PAP.

KW Tissue plasminogen activator-like protease; t-PALP; human;

thrombosis; peripheral arterial occlusion; pulmonary embolism;

XX

OS Homo sapiens.

XX

| | | |
|----|--|---|
| FH | Key | Location/Qualifiers |
| FT | Peptide | 1..21 /label= Sig_peptide |
| FT | Protein | 22..263 /label= Mat_protein |
| FT | Domain | 25..84 /note= "kringle domain" |
| FT | Domain | 85..263 /note= "protease domain" |
| FT | Peptide | 22..31 /note= "epitope-bearing region" |
| FT | Peptide | 35..44 /note= "epitope-bearing region" |
| FT | Peptide | 71..81 /note= "epitope-bearing region" |
| FT | Peptide | 91..107 /note= "epitope-bearing region" |
| FT | Peptide | 119..128 /note= "epitope-bearing region" |
| FT | Peptide | 138..147 /note= "epitope-bearing region" |
| FT | Peptide | 155..167 /note= "epitope-bearing region" |
| FT | Peptide | 193..203 /note= "epitope-bearing region" |
| FT | Peptide | 206..215 /note= "epitope-bearing region" |
| FT | Peptide | 227..237 /note= "epitope-bearing region" |
| FT | Peptide | 243..252 /note= "epitope-bearing region" |
| XX | | |
| PN | WO854199-A1. | |
| PD | 03-DEC-1998. | |
| PF | 27-MAY-1998; 98WO-US10728. | |
| PR | 28-MAY-1997; 97US-0048000. | |
| PA | (HUMA-) HUMAN GENOME SCI INC. | |
| PI | Edner R, Moore PA, Ruben SM; | |
| DR | WPI; 1999-070207/06. | |
| XX | N-PSDB; AAV99636. | |
| PT | New tissue plasminogen activator-like protease - useful in the diagnosis and treatment of circulatory system-related disorders | |
| XX | Claim 1; Page 56-57; 76pp; English. | |
| XX | This is the amino acid sequence of tissue plasminogen activator-like protease (t-PALP), a novel member of the serine protease family that shares sequence homology to human tissue plasminogen activator (see AA087770). The t-PALP sequence was deduced from a cDNA clone (see AAV99636) derived from activated monocytes. The 2.5 kb t-PALP message has also been detected in heart, brain, lung, placenta, liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate, testis, ovary, small intestine, colon and peripheral blood leukocytes. Isolated nucleic acids encoding amino acids -21 to 242, -20 to 242, 1-242, 4-63 (kringle domain) and 64-242 (protease domain) of t-PALP, or encoding epitope-bearing portions of t-PALP, are also claimed, as are recombinant vectors, host cells, and methods for producing t-PALP polypeptides. t-PALP may be used to detect and treat disorders related to the circulatory system, and to identify agonists and antagonists of t-PALP activity. The homology between t-PALP and tPA indicates that t-PALP may be involved in the regulation of normal and abnormal clotting in e.g. stroke, deep-vein thrombosis, peripheral arterial occlusion, pulmonary embolism and myocardiothrombosis. | |
| XQ | Sequence 263 AA: | |

arterial occlusion, blood coagulation disorders, (auto)immune system disorders e.g. human immunodeficiency syndrome, rheumatoid arthritis, graft-versus-host disease, thyroiditis, insulin dependent diabetes and inflammatory eye disease, allergic reactions e.g. asthma, cardiovascular diseases e.g. heart disease, arrhythmia and myocardial ischaemia, hyperproliferative disorders, cancers, hypertrophic scars and keloids, neurological diseases e.g. Creutzfeldt-Jakob syndrome, neurodegenerative disorders e.g. Alzheimer's disease and Parkinson's disease and infectious disease e.g. viral, bacterial and fungal infections. The t-PALP sequences are also useful for drug screening. The t-PALP nucleotides are useful as chromosome markers and are involved in gene therapy.

Sequence 263 AA;

Query Match 100.0%; Score 1289; DB 22; Length 263;
Best Local Similarity 100.0%; Pred. No. 4.6e-114;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFMNGHLHYREDQTSAPGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGRGW 60
DB 22 SGGCFMNGHLHYREDQTSAPGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGRGW 81
QY 61 CVYSGAGVPEKRPCEDLRCPEPTTSQALPAFTTEIOEASGPGADVQYFAPANALPARS 120
DB 82 CVYSGAGVPEKRPCEDLRCPEPTTSQALPAFTTEIOEASGPGADVQYFAPANALPARS 141
QY 121 EAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMWWIIAIGAGIILIGYSYKRGKDLK 180
DB 142 EAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMWWIIAIGAGIILIGYSYKRGKDLK 201
QY 181 EOHDOKVCEREMQRTLPISAFNPTCEIYDEKTVVHTSQTVPVPOEGSTPLMGAGATP 240
DB 202 EOHDOKVCEREMQRTLPISAFNPTCEIYDEKTVVHTSQTVPVPOEGSTPLMGAGATP 261
QY 241 GA 242
DB 262 GA 263

RESULT 4
AA086149
ID AA086149 standard; Protein; 263 AA.

AC AA086149;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human PRO264 polypeptide.
XX
XX Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;
KW leukaemia; neuronal disorder; stromal disorder; blastocoealic disorder;
KW inflammatory disorder; immune disorder; angiogenic disorder;
KW cytostatic; neuroprotective.
XX
OS Homo sapiens.
XX
PN W0200153466-AA1.
XX
PD 26-JUL-2001.
XX
PF 11-FEB-2000; 2000WO-US03565.
XX
PR 08-MAR-1999; 99WO-US05028.
PR 11-MAR-1999; 99US-123972P.
PR 11-MAY-1999; 99US-133459P.
PR 02-JUN-1999; 99WO-US12252.
PR 22-JUN-1999; 99US-140650P.
PR 22-JUN-1999; 99US-140653P.
PR 20-JUL-1999; 99US-144758P.
PR 26-JUL-1999; 99US-145698P.
PR 28-JUL-1999; 99US-146222P.
PR 17-AUG-1999; 99US-149395P.
PR 31-AUG-1999; 99US-151689P.

PR 01-SEP-1999; 99WO-US02011.
PR 15-SEP-1999; 99WO-US21090.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 05-JAN-2000; 2000WO-US00219.
XX

(GENTH) GENENTECH INC.

PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
PI Masters SA, Pan J, Pilti RM, Roy MA, Smith V, Stone DM;
PI Matanabe CK, Wood WI;
XX

DR WPI; 2002-205567/26.

DR N-PSDB; ABR40275.

PT Thirty five nucleic acids encoding PRO polypeptides, useful for
PT treating benign or malignant tumours, leukaemias and lymphoid
PT malignancies, inflammatory, angiogenic and immunologic disorders -
XX
PS Claim 61, Fig 44; 302pp; English.

XX The present invention relates to the isolation of novel human PRO
CC polypeptides and the polynucleotide sequences encoding them. The
CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are
CC useful for treating benign or malignant tumours (e.g. renal, kidney,
CC bladder, breast, etc), leukaemias and lymphoid malignancies, other
CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,
CC macrophagal, stromal and blastocoealic disorders, inflammatory, immune
CC and angiogenic disorders. The polynucleotide sequences are also
CC useful in gene therapy. AA086128-AA086162 represent the human PRO
CC polypeptides of the invention.

Sequence 263 AA;

Query Match 99.8%; Score 1286; DB 23; Length 263;
Best Local Similarity 99.6%; Pred. No. 8.8e-114;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFMNGHLHYREDQTSAPGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGRGW 60
DB 22 SGGCFMNGHLHYREDQTSAPGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGRGW 81
QY 61 CVYSGAGVPEKRPCEDLRCPEPTTSQALPAFTTEIOEASGPGADVQYFAPANALPARS 120
DB 82 CVYSGAGVPEKRPCEDLRCPEPTTSQALPAFTTEIOEASGPGADVQYFAPANALPARS 141
QY 121 EAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMWWIIAIGAGIILIGYSYKRGKDLK 180
DB 142 EAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMWWIIAIGAGIILIGYSYKRGKDLK 201
QY 181 EOHDOKVCEREMQRTLPISAFNPTCEIYDEKTVVHTSQTVPVPOEGSTPLMGAGATP 240
DB 202 EOHDOKVCEREMQRTLPISAFNPTCEIYDEKTVVHTSQTVPVPOEGSTPLMGAGATP 261
QY 241 GA 242
DB 262 GA 263

RESULT 5
AAB43237
ID AAB43237 standard; Protein; 263 AA.

AC AAB43237;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF3001 polypeptide sequence SEQ ID NO:6002.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antiparkinsonian; nocotropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;

immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antineoplastic; antiviral; antibacterial; antifungal; antineoplastic; antihypertensive; antianemic; gene therapy; cancer; proliferative disorder; hyperextension; neurodegenerative disorder; osteoarthritis; graft vs host disease; cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS; cholesterol ester storage; systemic lupus erythematosus; infection; severe combined immunodeficiency; malaria; autoimmune disorder; asthma; allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; bone damage; cartilage damage; antiinflammatory disease; coagulation; thrombosis; contraceptive.

OS Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CUBA-) CUBAGEN CORP.

XX Shinketsu RA, Leach M;

XX WPI; 2000-602362/57.

XX N-PSDB; AAC77446.

XX Novel nucleic acids and peptides derived from open reading frame X, useful for treating e.g. cancers, proliferative disorders, neurodegenerative disorders and cardiovascular disease -

XX Claim 11; Page 5181-5182; 5507pp; English.

CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397, which represent the human ORFX open reading frames 1 to 3161. The ORFX sequences have activities such as: cytoskeletal; hepatotropic; vulnary; antiproliferative; antiparkinsonian; nootropic; neuroprotective; osteoprotective; anticonvulsant; antirheumatic; immunosuppressant; immunostimulant; cardiac; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antineoplastic; antibacterial; antiviral; antifungal; antirheumatic; antihypertensive; antianemic. The sequences can be used for determining the presence of or predisposition to or preventing or treating pathological conditions associated with an ORFX-associated disorder. The nucleic acids can be used to express ORFX proteins in gene therapy vectors. The proteins and nucleic acids may be used to treat cancers, proliferative disorders, neurodegenerative disorders, osteoarthritis, graft vs host disease, cardiovascular disease, diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage, systemic lupus erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to enhance coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 263 AA;

SO

Query Match 99.3%; Score 1280; DB 21; Length 263;

Best Local Similarity 99.2%; Pred. No. 3.3e-113;

Matches 240; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGGCFWMDNGHLYREDOTSPAPGLRCLNWLDAOSGLASAPVSGAGNHSCRNPDDEPRGPW 60

DB 22 SGGCFWMDNGHLYREDOTSPAPGLRCLNWLDAOSGLASAPVSGAGNHSCRNPDDEPRGPW 81

QY 61 CVSGAGVPEKRPCEDELRCPETTSQALPAFTTETQESSEPGADEVVPAPNALPARS 120

DB 82 CVSGAGVPEKRPCEDELRCPETTSQALPAFTTETQESSEPGADEVVPAPNALPARS 141

QY 121 EAAAVQPIGISOVRMNSKEKDLGTGLVYAGITMMVITIAAGIILGYSYRGKDL 180

DB 142 EAAVQPIGISOVRMNSKEKDLGTGLVYAGITMMVITIAAGIILGYSYRGKDL 201

QY 181 EQHDQKVEREMQRTTLPISAFNPTCEIVDEKTVVHTSQTPVDEGSGTPLMGAGTP 240

DB 202 EQHDQKVEREMQRTTLPISAFNPTCEIVDEKTVVHTSQTPVDEGSGTPLMGAGTP 261

QY 241 GA 242

DB 262 GA 263

RESULT 6

AA093748

ID AA093748 standard; Protein; 263 AA.

AC AA093748;

DT 06-NOV-2001 (first entry)

DE Human polypeptide, SEQ ID NO: 3727.

XX Human, full length cDNA; cDNA synthesis; oligo-capping.

XX Homo sapiens.

XX EPI130094-A2.

XX 05-SEP-2001.

XX 07-JUL-2000; 2000EP-0114089.

XX 08-JUL-1999; 99EP-0194486.

XX 11-JAN-2000; 2000JP-0118774.

XX 02-MAY-2000; 2000JP-0183765.

XX (HELI-) HELIX RES INST.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2001-524255/58.

XX N-PSDB; AAK94700.

PT 830 Primers useful for synthesizing full length cDNA clones and their use in genetic manipulation -

XX Claim 8; SEQ ID NO 3727; 1380pp + sequence listing; English.

XX The invention relates to primers for synthesizing full length cDNA clones. 830 cDNA molecules encoding a human protein have been isolated and nucleotide sequences of 5' - and 3' -ends of the cDNA molecules have been determined. Primers for synthesizing the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesized by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence is a polypeptide encoded by a full length human cDNA of the invention.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 263 AA;

SO

Query Match 98.9%; Score 1275; DB 22; Length 263;

Best Local Similarity 99.2%; Pred. No. 9.8e-113;

Matches 240; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGGCFWMDNGHLYREDOTSPAPGLRCLNWLDAOSGLASAPVSGAGNHSCRNPDDEPRGPW 60

DB 22 SGGCFWMDNGHLYREDOTSPAPGLRCLNWLDAOSGLASAPVSGAGNHSCRNPDDEPRGPW 81

QY 61 CYSGEAGVPEKRPCEDLRCPEITTSQALPAFTTEIOEASBEPGADDEVQVFPANALPARS 120
 DB 82 CYSGEAGVPEKRPCEGRCEITTSQALPAFTTEIOEASBEPGADDEVQVFPANALPARS 141
 QY 121 EAAAVQPIVIGISQVRNMSKEKKDGLGTGLVGLITMWVIIAIGAGIILGYSYRGKDLK 180
 DB 142 EAAAVQPIVIGISQVRNMSKEKKDGLGTGLVGLITMWVIIAIGAGIILGYSYRGKDLK 201
 QY 181 EOHDKVCEREMQRIITLPLSAFTNPTCEIVDEKTVVHTSQTVPDPOEGSTPLMGQAGTP 240
 DB 202 EOHDKVCEREMQRIITLPLSAFTNPTCEIVDEKTVVHTSQTVPDPOEGSTPLMGQAGTP 261
 QY 241 GA 242
 DB 262 GA 263

RESULT 7
 ID AAY05220 standard; Protein; 286 AA.
 AC AAY05220;
 DT 17-JUN-1999 (first entry)
 DE Kringle1 protein sequence.
 XX Kringle1; cancer; inflammation; autoimmunity; rheumatoid arthritis; AIDS;
 KM CNS inflammation; cerebellar degeneration; Alzheimer's disease; asthma;
 KM Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
 KM neurological abnormality; ischemia reperfusion injury; ischaemic injury;
 KM cardiovascular disease; kidney disease; liver disease; aplastic anaemia;
 KM myocardial infarction; hypotension; hypertension; allergy; infection;
 KM myelodysplastic syndrome; haematologic abnormality; diagnosis; therapy;
 KM male pattern baldness.
 XX Homo sapiens.
 OS WO9911788-A1.
 PN 11-MAR-1999.
 PD 02-SEP-1998; 98WO-US18270.
 PF 02-SEP-1998; 98US-0144889.
 PR 01-SEP-1998; 97US-0056032.
 PR 02-SEP-1997; 97US-0056032.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA Albione EF, Kikly KK;
 PI WPI; 1999-214707/18.
 DR N-PSDB; AAX28355.
 XX New kringle1 polypeptides and polynucleotides
 PT Claim 14; Page 33; 42pp; English.

This sequence is a Kringle1 polypeptide of the invention.
 CC The kringle1 polypeptides (I) are used to screen for agonists and
 CC antagonists. Agonists are used to treat subjects in need of enhanced
 CC activity or expression of (I). Antagonists are used to treat subjects
 CC having need to inhibit the activity or expression of (I). The methods can
 CC be used to treat conditions such as cancer, inflammation, autoimmunity,
 CC allergy, asthma, rheumatoid arthritis, CNS inflammation, cerebellar
 CC degeneration, Alzheimer's disease, Parkinson's disease, multiple
 CC sclerosis, amyotrophic lateral sclerosis, head injury, damage and other
 CC neurological abnormalities, ischemia reperfusion injury, cardiovascular
 CC disease, kidney disease, liver disease, ischaemic injury, myocardial
 CC infarction, hypotension, hypertension, AIDS, myelodysplastic syndromes
 CC and other haematologic abnormalities, aplastic anaemia, male pattern
 CC baldness, and bacterial, fungal, protozoan and viral infections. The

CC Kringle1 polypeptides may also be used to generate antibodies.
 CC Determining the presence or absence of mutations in, and analyzing for
 CC the presence or absence of expression of, Kringle1 polynucleotides can be
 CC used to diagnose a disease or susceptibility to a disease related to
 CC expression or activity of Kringle1 protein. The polynucleotides may also
 CC be used for chromosome identification, and mapping.

SO Sequence 286 AA;
 Query Match 82.5%; Score 1063.5; DB 20; Length 286;
 Best Local Similarity 86.2%; Pred. No. 1.3e-92;
 Matches 206; Conservative 5; Mismatches 27; Indels 1; Gaps 1;

QY 1 SGGCFMNDGHLVREDOTSPAPGRCLNMDAOSGLASAPVSGAGNHSYCRNDEDPGRPW 60
 DB 22 SGGCFMNDGHLVREDOTSPAPGRCLNMDAOSGLASAPVSGAGNHSYCRNDEDPGRPW 81
 QY 61 CYSGEAGVPEKRPCEDLRCPEITTSQALPAFTTEIOEASBEPGADDEVQVFPANALPAR 119
 DB 82 CYSGEAGVPEKRPCEDLRCPEITTSQALPAFTTEIOEASBEPGADDEVQVFPANALPAR 141
 QY 120 SEAAVQPIVIGISQVRNMSKEKKDGLGTGLVGLITMWVIIAIGAGIILGYSYRGKDL 179
 DB 142 SEAAVQPIVIGISQVRNMSKEKKDGLGTGLVGLITMWVIIAIGAGIILGYSYRGKDL 201
 QY 180 EOHDKVCEREMQRIITLPLSAFTNPTCEIVDEKTVVHTSQTVPDPOEGSTPLMGQAG 238
 DB 202 EOHDKVCEREMQRIITLPLSAFTNPTCEIVDEKTVVHTSQTVPDPOEGSTPLMGQAG 260

RESULT 8
 ID ABB37905 standard; Peptide; 66 AA.
 AC ABB37905;
 DT 04-FEB-2002 (first entry)
 DE Peptide #5411 encoded by human foetal liver single exon probe.
 XX Human; foetal liver; gene expression; single exon nucleic acid probe.
 KM Homo sapiens.
 OS WO200157277-A2.
 PN 30-JAN-2001; 2001WO-US00669.
 PF 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-483447/52.
 DR Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human foetal liver -
 XX Claim 27; SEQ ID NO 30540; 639pp + sequence listing; English.

The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human

CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 66 AA;

Query Match 25.0%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 4,4e-23;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 TTSQLPAPFTTIOGASGPGADEVOVFAPANALPARSEAAAVQPVIGISQRYMNSKEK 142
Db 1 TTSQLPAPFTTIOGASGPGADEVOVFAPANALPARSEAAAVQPVIGISQRYMNSKEK 60
Qy 143 KDLGTL 148
Db 61 KDLGTL 66

RESULT 9
ABR23159
ID ABR23159 standard; Protein; 66 AA.
XX
AC ABR23159;
XX
DT 23-JUN-2002 (first entry)
XX
DE Protein #5158 encoded by probe for measuring heart cell gene expression.
XX
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN MO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JUN-2001; 2001MO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
XX
PT hearts -
XX
PS Claim 15; SEQ ID No 24929; 530bp; English.
XX
XX

CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA2153-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting, the
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 66 AA;

Query Match 25.0%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 4,4e-23;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 TTSQLPAPFTTIOGASGPGADEVOVFAPANALPARSEAAAVQPVIGISQRYMNSKEK 142
Db 1 TTSQLPAPFTTIOGASGPGADEVOVFAPANALPARSEAAAVQPVIGISQRYMNSKEK 60
Qy 143 KDLGTL 148
Db 61 KDLGTL 66

RESULT 10
AAM58537
ID AAM58537 standard; Protein; 66 AA.
XX
AC AAM58537;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30642.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN MO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JUN-2001; 2001MO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
XX
PT brains -
XX
PS Example 4; SEQ ID NO: 30642; 650bp + Sequence Listing; English.
XX
XX

CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
XX

Qy 83 TTSQLPAPFTTIOGASGPGADEVOVFAPANALPARSEAAAVQPVIGISQRYMNSKEK 142

Query Match 25.0%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 4,4e-23;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db      1  TTSQLPATTETIOASBGPAGDEVQVFAPANALPARSEAAAVOPVIGISQVRNMSKX 60
QY      143 KDLGTL 148
XX      |||||
DT      61 KDLGTL 66
Db      61 KDLGTL 66

RESULT 11
AAM1037
ID      AAM1037 standard; Protein; 66 AA.
XX
AC      AAM1037;
XX
DT      06-NOV-2001 (first entry)
XX
DE      Human bone marrow expressed probe encoded protein SEQ ID NO: 31343.
XX
KW      Human; bone marrow expressed exon; gene expression analysis; probe;
KW      microarray; cancer; leukemia; lymphoma; myeloma.
XX
OS      Homo sapiens.
XX
PN      WO200157276-A2.
XX
PD      09-AUG-2001.
XX
PF      30-JAN-2001; 2001WO-US00668.
XX
PI      04-FEB-2000; 2000US-0180312.
XX
PR      26-MAY-2000; 2000US-0207456.
XX
PR      30-JUN-2000; 2000US-0608408.
XX
PR      03-AUG-2000; 2000US-0632366.
XX
PR      21-SEP-2000; 2000US-0234687.
XX
PR      27-SEP-2000; 2000US-0236359.
XX
PR      04-OCT-2000; 2000GB-0024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR      WPI; 2001-488900/53.
XX
PT      Human genome-derived single exon nucleic acid probes useful for
XX      analyzing gene expression in human bone marrow -
XX
PS      Example 4; SEQ ID NO: 31343; 658bp + Sequence listing; English.
XX
CC      The present invention provides a number of single exon nucleic acid
XX      probes which are derived from genomic sequences expressed in the human
XX      bone marrow. They can be used to measure gene expression in bone marrow
XX      samples, which may enable the improved diagnosis and treatment of cancers
XX      such as lymphoma, leukemia and myeloma. The present sequence is a
XX      protein encoded by one of the probes of the invention.
XX
SQ      Sequence 66 AA;

Query Match      25.0%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 4.4e-23;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      83  TTSQLPATTETIOASBGPAGDEVQVFAPANALPARSEAAAVOPVIGISQVRNMSKX 142
Db      1  TTSQLPATTETIOASBGPAGDEVQVFAPANALPARSEAAAVOPVIGISQVRNMSKX 60
QY      143 KDLGTL 148
Db      61 KDLGTL 66

RESULT 12,
AAM1800
ID      AAM1800 standard; Protein; 66 AA.

```

```

XX      AAM1800;
AC      12-OCT-2001 (first entry)
XX
DT      Peptide #5234 encoded by probe for measuring cervical gene expression.
XX
DE      Probe; human; microarray; gene expression; cervical epithelial cell;
XX      cervical cancer.
XX
KW      Homo sapiens.
XX
PN      WO200157278-A2.
XX
PD      09-AUG-2001.
XX
PF      30-JAN-2001; 2001WO-US00670.
XX
PI      04-FEB-2000; 2000US-0180312.
XX
PR      26-MAY-2000; 2000US-0207456.
XX
PR      30-JUN-2000; 2000US-0608408.
XX
PR      03-AUG-2000; 2000US-0632366.
XX
PR      21-SEP-2000; 2000US-0234687.
XX
PR      27-SEP-2000; 2000US-0236359.
XX
PR      04-OCT-2000; 2000GB-0024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR      WPI; 2001-488901/53.
XX
PT      Human genome-derived single exon nucleic acid probes useful for
XX      analyzing gene expression in human cervical epithelial cells -
XX
PS      Claim 27; SEQ ID No 23626; 487bp; English.
XX
CC      The present invention relates to human single exon nucleic acid probes
XX      (SENP; see A110068-A1126459). The present sequence is a peptide encoded
XX      by one such probe. The SENPs are derived from human HeLa cells. The SENPs
XX      can be used to produce a single exon microarray, which can be used for
XX      measuring human gene expression in a sample derived from human cervical
XX      epithelial cells. By measuring gene expression, the probes are therefore
XX      useful in grading and/or staging of diseases of the cervix, notably
XX      cervical cancer.
XX
CC      Note: The sequence data for this patent did not form part of the printed
XX      specification, but was obtained in electronic format directly from WIPO
XX      at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ      Sequence 66 AA;

Query Match      25.0%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 4.4e-23;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      83  TTSQLPATTETIOASBGPAGDEVQVFAPANALPARSEAAAVOPVIGISQVRNMSKX 142
Db      1  TTSQLPATTETIOASBGPAGDEVQVFAPANALPARSEAAAVOPVIGISQVRNMSKX 60
QY      143 KDLGTL 148
Db      61 KDLGTL 66

RESULT 13
AAM31314
ID      AAM31314 standard; Protein; 66 AA.
XX
AC      AAM31314;
XX
DT      17-OCT-2001 (first entry)
XX
DE      Peptide #5351 encoded by probe for measuring placental gene expression.

```

XX Probe; microarray; human; placenta; antenatal diagnosis;
KM genetic disorder.
XX
XX Homo sapiens.
OS
XX WO200157272-A2.
PN
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00663.
XX
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
XX Claim 27; SEQ ID No 31583; 654bp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP;
CC see AAI31315-AA157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta; The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
XX Sequence 66 AA;
SQ
Query Match 25.0%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 4,4e-23;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 83 TTSQLPAPFTTETIOFASGPGADDEVOPAPANNALPASEEAAVQVITGISQVRNMSKEX 142
DB 1 TTSQLPAPFTTETIOFASGPGADDEVOPAPANNALPASEEAAVQVITGISQVRNMSKEX 60
QY 143 KDLGTL 148
DB 61 KDLGTL 66
RESULT 14
ABG40828
ID ABG40828 standard; Peptide; 66 AA.
AC
XX ABG40828;
XX
XX 19-AUG-2002 (first entry)
XX
XX Human peptide encoded by genome-derived single exon probe SEQ ID 30493.
DE
XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
KM chronic obstructive pulmonary disease; interstitial lung disease;
KM familial idiopathic pulmonary fibrosis; neurofibromatosis;
KM tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
KM Heremansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KM pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
KM pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KM primary ciliary dyskinesia; pulmonary hypertension;
KM hyaline membrane disease.
XX

OS Homo sapiens.
XX
XX WO200186003-A2.
PN
XX 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US00665.
XX
XX 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -
XX
XX Claim 27; SEQ ID No 30493; 634bp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridise at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung; comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberosus sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Heremansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a peptide/protein
CC encoded by a single exon probe of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 66 AA;
SQ
Query Match 25.0%; Score 322; DB 23; Length 66;
Best Local Similarity 100.0%; Pred. No. 4,4e-23;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 TTSGALPATTETIQASRGPGADEVQVAPAPNALPASEMAAVQVITGISQVRNMSKEK 142
 DB 1 TTSGALPATTETIQASRGPGADEVQVAPAPNALPASEMAAVQVITGISQVRNMSKEK 60
 QY 143 KDLGTL 148
 DB 61 KDLGTL 66

RESULT 15

AA12615
 ID AAY12615 standard; Protein; 56 AA.

XX AAY12615;
 AC AAY12615;

DT 22-JUN-1999 (first entry)
 DE Human 5' EST secreted protein SEQ ID NO: 280 from WO 9906553.

XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide;
 KW upstream regulatory sequence; cytokine activity; cell proliferation;
 KW differentiation; haematopoiesis regulation; tissue growth regulation;
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; antiinflammatory; tumour inhibition; antitumour.

XX Homo sapiens.
 OS

XX MO9906553-A2.
 PN

XX 11-FEB-1999.
 PD

XX 31-JUL-1998; 98WO-IB01237.
 PF

XX 01-AUG-1997; 97US-0905051.
 PR

XX (GSEST) GENSET.
 PA

XX DucJert A, Dumas Milne Edwards J, Lacroix B;
 PI

XX WPI; 1999-153783/13.
 DR

XX N-PSDB; AAX41473.
 DR

XX New nucleic acids encoding human secreted proteins - obtained from
 PT CDNA libraries derived from umbilical cord, lymph ganglia,
 PT lymphocytes and placental tissue

XX Claim 34; Page 376; 411pp; English.
 PS

XX AAX41379 to AAX41526 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAY12521 to
 CC AAY12668, respectively. The proteins given represent the signal peptide
 CC and an N-terminal fragment of a secreted protein. The nucleic acid
 CC sequences can be used for producing secreted human gene products. They
 CC can also be used to develop products for diagnosis and therapy. The
 CC proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, haematopoiesis regulating
 CC activity, tissue growth regulation activity, reproductive hormone
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
 CC thrombolytic activity, receptor/ligand activity, antiinflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.

XX Sequence 56 AA;
 SQ

Query Match 15.7%; Score 202; DB 20; Length 56;
 Best Local Similarity 97.1%; Pred. No. 8.7e-12;
 Matches 34; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFWNGHLYREDQTSAPGLRCINWLDAGSL 35
 DB 22 SGGCFWNGHLYREDQTSAPGLRCINWLDAGSL 56

Search completed: April 7, 2003, 09:18:32
 Job time : 45.2807 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:19:28 ; Search time 17.106 Seconds

(without alignments)
864,895 Million cell updates/sec

Title: US-10-057-951-2_COPY_22_263

Perfect score: 1289
Sequence: 1 SGGCFWMDNGHLYREDQTSFPA.....PVDPQEGSTPLMGAGTPGA 242

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1 | 1289 | 100.0 | 263 | 9 US-10-102-704-2 | Sequence 2, Appl1 |
| 2 | 1289 | 100.0 | 263 | 9 US-10-057-951-2 | Sequence 2, Appl1 |
| 3 | 1289 | 100.0 | 263 | 10 US-09-084-491A-2 | Sequence 2, Appl1 |
| 4 | 1289 | 100.0 | 263 | 10 US-09-084-491A-2 | Sequence 2, Appl1 |
| 5 | 154.5 | 12.0 | 527 | 9 US-09-987-457-18 | Sequence 18, Appl1 |
| 6 | 154.5 | 12.0 | 527 | 9 US-09-987-457-18 | Sequence 18, Appl1 |
| 7 | 154.5 | 12.0 | 527 | 9 US-09-987-457-18 | Sequence 18, Appl1 |
| 8 | 154.5 | 12.0 | 527 | 9 US-09-987-457-18 | Sequence 18, Appl1 |
| 9 | 142 | 11.0 | 160 | 9 US-09-335-325-35 | Sequence 35, Appl1 |
| 10 | 142 | 11.0 | 160 | 9 US-09-335-325-35 | Sequence 35, Appl1 |
| 11 | 142 | 11.0 | 160 | 9 US-09-335-325-35 | Sequence 35, Appl1 |
| 12 | 142 | 11.0 | 160 | 9 US-09-335-325-35 | Sequence 35, Appl1 |
| 13 | 142 | 11.0 | 250 | 9 US-10-131-241-30 | Sequence 30, Appl1 |
| 14 | 142 | 11.0 | 250 | 9 US-10-131-241-30 | Sequence 30, Appl1 |
| 15 | 142 | 11.0 | 250 | 9 US-10-131-241-30 | Sequence 30, Appl1 |
| 16 | 142 | 11.0 | 250 | 9 US-10-131-241-30 | Sequence 30, Appl1 |
| 17 | 142 | 11.0 | 339 | 9 US-09-335-325-35 | Sequence 3, Appl1 |
| 18 | 142 | 11.0 | 339 | 9 US-10-131-241-30 | Sequence 3, Appl1 |
| 19 | 142 | 11.0 | 339 | 9 US-09-761-120-3 | Sequence 3, Appl1 |

| | | | | | |
|----|-------|------|-----|---------------------|--------------------|
| 20 | 142 | 11.0 | 352 | 9 US-09-335-325-40 | Sequence 40, Appl1 |
| 21 | 142 | 11.0 | 352 | 9 US-10-131-241-40 | Sequence 40, Appl1 |
| 22 | 142 | 11.0 | 352 | 10 US-09-761-120-40 | Sequence 40, Appl1 |
| 23 | 142 | 11.0 | 352 | 10 US-09-761-120-40 | Sequence 40, Appl1 |
| 24 | 142 | 11.0 | 378 | 9 US-09-335-325-42 | Sequence 42, Appl1 |
| 25 | 142 | 11.0 | 378 | 9 US-10-131-241-42 | Sequence 42, Appl1 |
| 26 | 142 | 11.0 | 458 | 10 US-09-873-676-1 | Sequence 1, Appl1 |
| 27 | 142 | 11.0 | 458 | 10 US-09-873-676-1 | Sequence 1, Appl1 |
| 28 | 142 | 11.0 | 559 | 10 US-09-946-893-4 | Sequence 4, Appl1 |
| 29 | 142 | 11.0 | 559 | 10 US-09-946-893-4 | Sequence 4, Appl1 |
| 30 | 142 | 11.0 | 576 | 10 US-09-946-893-6 | Sequence 6, Appl1 |
| 31 | 142 | 11.0 | 791 | 9 US-09-967-386-1 | Sequence 1, Appl1 |
| 32 | 142 | 11.0 | 810 | 10 US-09-967-386-1 | Sequence 1, Appl1 |
| 33 | 138 | 10.7 | 79 | 9 US-09-335-325-8 | Sequence 8, Appl1 |
| 34 | 138 | 10.7 | 79 | 9 US-10-131-241-8 | Sequence 8, Appl1 |
| 35 | 138 | 10.7 | 79 | 10 US-09-753-064-2 | Sequence 2, Appl1 |
| 36 | 138 | 10.7 | 79 | 10 US-09-753-064-2 | Sequence 2, Appl1 |
| 37 | 132 | 10.2 | 411 | 9 US-09-880-503-3 | Sequence 3, Appl1 |
| 38 | 132 | 10.2 | 411 | 9 US-09-880-503-3 | Sequence 3, Appl1 |
| 39 | 132 | 10.2 | 431 | 9 US-10-076-421-2 | Sequence 1, Appl1 |
| 40 | 131.5 | 10.2 | 96 | 10 US-09-880-503-9 | Sequence 9, Appl1 |
| 41 | 131.5 | 10.2 | 143 | 10 US-09-880-503-8 | Sequence 8, Appl1 |
| 42 | 131.5 | 10.2 | 88 | 10 US-09-880-503-1 | Sequence 1, Appl1 |
| 43 | 131 | 10.2 | 135 | 10 US-09-880-503-4 | Sequence 4, Appl1 |
| 44 | 131 | 10.2 | 138 | 9 US-10-237-667-12 | Sequence 12, Appl1 |
| 45 | 131 | 10.2 | 138 | 9 US-10-237-708-12 | Sequence 12, Appl1 |

ALIGNMENTS

RESULT 1
US-10-102-704-2
Sequence 2, Application US/10102704
Patient No. US20020164768A1
GENERAL INFORMATION:
APPLICANT: Moore et al.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protein
FILE REFERENCE: P9378C1
CURRENT APPLICATION NUMBER: US/10/102,704
CURRENT FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/084,491
PRIOR FILING DATE: 1998-05-27
PRIOR APPLICATION NUMBER: 60/048,000
PRIOR FILING DATE: 1997-05-28
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-10-102-704-2

Query Match 100.0%; Score 1289; DB 9; Length 263;
Best Local Similarity 100.0%; Pred. No. 4e-110; Mismatches 0; Gaps 0;
Matches 242; Conservative 0; Indels 0;

| | | | | |
|----|-----|-------------------------|---|----------------------|
| Qy | 1 | SGGCFWMDNGHLYREDQTSFPA | GLRCINLWIDAGSGLASAPVSGAGNHSYCHNPDEDRGFW | 60 |
| Db | 22 | SGGCFWMDNGHLYREDQTSFPA | GLRCINLWIDAGSGLASAPVSGAGNHSYCHNPDEDRGFW | 81 |
| Qy | 61 | CYVSGAGVPEKPCEDRCPTTSOL | PAFTTEIOESSEGGADEVQVAPANALPARS | 120 |
| Db | 82 | CYVSGAGVPEKPCEDRCPTTSOL | PAFTTEIOESSEGGADEVQVAPANALPARS | 141 |
| Qy | 121 | EAAAVQPVGISQVRKMSKEXD | IGTGVVGIIMVYIIAIGAGIILGYSYKRRKDK | 180 |
| Db | 142 | EAAAVQPVGISQVRKMSKEXD | IGTGVVGIIMVYIIAIGAGIILGYSYKRRKDK | 201 |
| Qy | 181 | EHDQVCEREMQRTTLP | LSAFNPFCEIVDEKTVVHTSQT | PVDPEGSTPLMGAGTP 240 |
| Db | 202 | EHDQVCEREMQRTTLP | LSAFNPFCEIVDEKTVVHTSQT | PVDPEGSTPLMGAGTP 261 |

QY 241 GA 242
DB 262 GA 263

RESULT 2

US-10-057-951-2
; Sequence 2, Application US/10057951
; Patent No. US2002017213A1
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: P378P1
; CURRENT APPLICATION NUMBER: US/10/057,951
; PRIORITY FILING DATE: 1997-05-29
; PRIORITY FILING DATE: 1999-10-04
; PRIORITY FILING DATE: 1998-05-27
; PRIORITY FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent In Ver. 3.1
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-951-2

Query Match 100.0%; Score 1289; DB 9; Length 263;
Best Local Similarity 100.0%; Pred. No. 4e-110;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGCGFMDNCHLYREDQTSAPAGLCNMLDAOSGLASAPVSGAGNHSYCRNPDEDPGFW 60
DB 22 SGCGFMDNCHLYREDQTSAPAGLCNMLDAOSGLASAPVSGAGNHSYCRNPDEDPGFW 81
QY 61 CYSVSGAGVPEKRPCEIDRCPTTSQALPAFTTEIOEASBEGGADDEVQVFPANALPARS 120
DB 82 CYSVSGAGVPEKRPCEIDRCPTTSQALPAFTTEIOEASBEGGADDEVQVFPANALPARS 141
QY 121 EAAAVOPVIGISQVRNMSKEKDLGTLGVGLITMVTIIAIGAGIILGYSYRGKDLK 180
DB 142 EAAAVOPVIGISQVRNMSKEKDLGTLGVGLITMVTIIAIGAGIILGYSYRGKDLK 201
QY 181 EOHDOVCEREMQRTLPISAFNPTCEIVDEKTVVHTSQTPVDPQSGSTPLMGQAGTP 240
DB 202 EOHDOVCEREMQRTLPISAFNPTCEIVDEKTVVHTSQTPVDPQSGSTPLMGQAGTP 261
QY 241 GA 242
DB 262 GA 263

RESULT 3

US-09-084-491A-2
; Sequence 2, Application US/09084491A
; Patent No. US20020061576A1
; GENERAL INFORMATION:
; APPLICANT: MOORE, PAUL A.
; APPLICANT: RUBEN, STEVEN M.
; APPLICANT: BENNER, REINHARD
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/084,491A
FILING DATE: 27-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: P378
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-084-491A-2

Query Match 100.0%; Score 1289; DB 10; Length 263;
Best Local Similarity 100.0%; Pred. No. 4e-110;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGCGFMDNCHLYREDQTSAPAGLCNMLDAOSGLASAPVSGAGNHSYCRNPDEDPGFW 60
DB 22 SGCGFMDNCHLYREDQTSAPAGLCNMLDAOSGLASAPVSGAGNHSYCRNPDEDPGFW 81
QY 61 CYSVSGAGVPEKRPCEIDRCPTTSQALPAFTTEIOEASBEGGADDEVQVFPANALPARS 120
DB 82 CYSVSGAGVPEKRPCEIDRCPTTSQALPAFTTEIOEASBEGGADDEVQVFPANALPARS 141
QY 121 EAAAVOPVIGISQVRNMSKEKDLGTLGVGLITMVTIIAIGAGIILGYSYRGKDLK 180
DB 142 EAAAVOPVIGISQVRNMSKEKDLGTLGVGLITMVTIIAIGAGIILGYSYRGKDLK 201
QY 181 EOHDOVCEREMQRTLPISAFNPTCEIVDEKTVVHTSQTPVDPQSGSTPLMGQAGTP 240
DB 202 EOHDOVCEREMQRTLPISAFNPTCEIVDEKTVVHTSQTPVDPQSGSTPLMGQAGTP 261
QY 241 GA 242
DB 262 GA 263

RESULT 4

US-09-864-761-38457
; Sequence 38457, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIORITY FILING DATE: 2001-05-23
; PRIORITY FILING DATE: 2000-02-04
; PRIORITY FILING DATE: 2000-02-04
; PRIORITY FILING DATE: 2000-05-26
; PRIORITY FILING DATE: US 09/632,366
; PRIORITY FILING DATE: 2000-08-03
; PRIORITY FILING DATE: 2000-10-04
; PRIORITY FILING DATE: 2000-09-27
; PRIORITY FILING DATE: PCT/US01/00666

PRIOR APPLICATION NUMBER: 60/238,331
PRIOR FILING DATE: 2000-05-10
NUMBER OF SEQ ID NOS: 194
SOFTWARE: PERL Program
SEQ ID NO 145
LENGTH: 562
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CD1
US-09-974-298-145

Query Match 12.0%; Score 154.5; DB 9; Length 562;
Best Local Similarity 39.6%; Pred. No. 4,7e-06;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFWDNGHLYREDOTSPAPGLRCLNWLDAQSGIASAPVS-----GAGNHSYCRNPDE 54
DB 127 CYEDOGISYRGTSWTAESGAECTNW--NSSALAQKPYSGRRPDALRLGNGNHYCRNPDR 184

QY 55 DPGPMCVYSGEAGVPEKRPCEDLRCPEPTS 85
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213

RESULT 8
US-09-969-271-7
Sequence 7, Application US/09969271
Patent No. US20020098179A1
GENERAL INFORMATION:
APPLICANT: Pfizer Inc. (A1) designated States except GB and EP (GB);
APPLICANT: Pfizer Limited (GB and EP (GB) only)
TITLE OF INVENTION: Pharmaceutical Combinations
FILE REFERENCE: PCS1095IAPME
CURRENT APPLICATION NUMBER: US/09/969,271
CURRENT FILING DATE: 2001-10-01
PRIOR APPLICATION NUMBER: GB 0025473.0
PRIOR FILING DATE: 2000-10-17
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PaateSEO for Windows Version 4.0
SEQ ID NO: 7
LENGTH: 562
TYPE: PRT
ORGANISM: Homo sapiens
US-09-969-271-7

Query Match 12.0%; Score 154.5; DB 10; Length 562;
Best Local Similarity 39.6%; Pred. No. 4,7e-06;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFWDNGHLYREDOTSPAPGLRCLNWLDAQSGIASAPVS-----GAGNHSYCRNPDE 54
DB 127 CYEDOGISYRGTSWTAESGAECTNW--NSSALAQKPYSGRRPDALRLGNGNHYCRNPDR 184

QY 55 DPGPMCVYSGEAGVPEKRPCEDLRCPEPTS 85
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213

RESULT 9
US-09-935-325-35
Sequence 35, Application US/09335325
Patent No. US20020164717A1
GENERAL INFORMATION:
APPLICANT: Folkman, M. Judah
O'Reilly, Michael
Cao, Yihai
Sim, B. Kim Lee
TITLE OF INVENTION: Angiostatin Fragments and Method of Use
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSER: Jones & Aekew

STREET: 191 Peachtree Street, 37th Floor
City: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 09/335,325
FILING DATE: 17-Jun-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/612,788
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Warren, William J.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 05213-0126
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 160 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: K1-2
SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-335-325-35

Query Match 11.0%; Score 142; DB 9; Length 160;
Best Local Similarity 37.2%; Pred. No. 1.3e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 4 CFWDNGHLYREDOTSPAPGLRCLNWLDA-----QSGIASAPVSGAGNHSYCRNPDEDPG 58
DB 1 CKTGKGVKVRGTSKTKNGITCOXMSSTSPHRPSPATHPSEGI-BENYCRNPNDNDPOG 59

QY 59 PWCYVSGEAGVPEKRPCEDLRCPEPTS 82
DB 60 PWCYITD---PEKRYDYCDILECEB 81

RESULT 10
US-10-131-241-35
Sequence 35, Application US/10131241
Publication No. US20030012792A1
GENERAL INFORMATION:
APPLICANT: Holaday, John W.
Fotlier, Anne H.
TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
and Regulating Angiogenesis Using Cancer Markers
FILE REFERENCE: 05213-0344 43170-271565
CURRENT APPLICATION NUMBER: US/10/131,241
CURRENT FILING DATE: 2002-07-22
PRIOR APPLICATION NUMBER: US 09/413,049
PRIOR FILING DATE: 1999-10-06
PRIOR APPLICATION NUMBER: US 09/316,802
PRIOR FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: US 60/086,586
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 65

SOFTWARE: PatentIn version 3.1
SEQ ID NO 35
LENGTH: 160
TYPE: PRT
ORGANISM: Homo sapiens
US-10-131-241-35

Query Match 11.0%; Score 142; DB 9; Length 160;
Best Local Similarity 37.2%; Pred. No. 1.3e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

OY 4 CFMDNGHLYREDQTSAPARGLCLNWLDA-----QSGLASAPVSGAGNHSYCRNPDEDPRG 58
DB 1 CKTGNGKXNYRGMTSKTKNGITCKQMSSTSPHRFRSPATHPSEGL-ENNYCRNPDPNDQ 59
OY 59 PWCYVSGEAGVPEKR--PCEDLRCP 82
DB 60 PWCYTTD---PEKRYDYCDLCECE 81

RESULT 11
US-09-761-120-35
Sequence 35, Application US/09761120
Patent No. US20020037847A1
GENERAL INFORMATION:
APPLICANT: Polkman, M. Judah
TITLE OF INVENTION: Nucleic Acids Encoding K-tingle 1-5 Region Fragments of Plasminogen
FILE REFERENCE: 05940-0151 (43171-252068)
CURRENT APPLICATION NUMBER: US/09/761,120
CURRENT FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 09/309,821
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 08/866,735
PRIOR FILING DATE: 1997-05-30
NUMBER OF SEQ ID NOS: 47
SOFTWARE: PatentIn version 3.0
SEQ ID NO 35
LENGTH: 160
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: K-tingle 1-2
US-09-761-120-35

Query Match 11.0%; Score 142; DB 10; Length 160;
Best Local Similarity 37.2%; Pred. No. 1.3e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

OY 4 CFMDNGHLYREDQTSAPARGLCLNWLDA-----QSGLASAPVSGAGNHSYCRNPDEDPRG 58
DB 1 CKTGNGKXNYRGMTSKTKNGITCKQMSSTSPHRFRSPATHPSEGL-ENNYCRNPDPNDQ 59
OY 59 PWCYVSGEAGVPEKR--PCEDLRCP 82
DB 60 PWCYTTD---PEKRYDYCDLCECE 81

RESULT 12
US-09-335-325-30
Sequence 30, Application US/09335325
Patent No. US20020164717A1
GENERAL INFORMATION:
APPLICANT: Polkman, M. Judah
O'Reilly, Michael
Cao, Yihai
Sim, B. Kim Lee
NUMBER OF INVENTION: Angiostatin Fragments and Method of Use
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones & Askew
STREET: 191 Peachtree Street, 37th Floor

CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/335,325
FILING DATE: 17-Jun-1999
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/612,788
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Warren, William U.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 05213-0126
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 250 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEetical: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: K1-3
SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-09-335-325-30

Query Match 11.0%; Score 142; DB 9; Length 250;
Best Local Similarity 37.2%; Pred. No. 2.3e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

OY 4 CFMDNGHLYREDQTSAPARGLCLNWLDA-----QSGLASAPVSGAGNHSYCRNPDEDPRG 58
DB 1 CKTGNGKXNYRGMTSKTKNGITCKQMSSTSPHRFRSPATHPSEGL-ENNYCRNPDPNDQ 59
OY 59 PWCYVSGEAGVPEKR--PCEDLRCP 82
DB 60 PWCYTTD---PEKRYDYCDLCECE 81

RESULT 13
US-10-131-241-30
Sequence 30, Application US/10131241
Publication No. US20030012792A1
GENERAL INFORMATION:
APPLICANT: Holaday, John W.
TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
FILE REFERENCE: 05213-0344 43170-271565
CURRENT APPLICATION NUMBER: US/10/131,241
CURRENT FILING DATE: 2002-07-22
PRIOR APPLICATION NUMBER: US 09/413,049
PRIOR FILING DATE: 1999-10-06
PRIOR APPLICATION NUMBER: US 09/316,802
PRIOR FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: US 60/086,586
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.1

Mon Apr 7 10:22:01 2003

us-10-057-951-2_copy_22_263.rapb

Page 6

SEQ ID NO 30
LENGTH: 250
TYPE: PRT
ORGANISM: Homo sapiens
US-10-131-241-30

Query Match 11.0%; Score 142; DB 9; Length 250;
Best Local Similarity 37.2%; Pred. No. 2.3e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 4 CFWDNGHLYREDQTSAPAGLRCLNWLDA-----QGLASAPVSGAGNHSYCRNPDEDPFG 58
DB 1 CKTGNGKNYRGTMSTKXGJTCOKMSSTSPHRPSPATHPSEGL-ENYCRNPDPDPOG 59
QY 59 PWCYVSGEAGVPEKR--PCEDLRCEP 82
DB 60 PWCYTTD---PEKRYDYCDILECEE 81

RESULT 14

US-09-761-120-30
Sequence 30 Application US/09761120
Patent No. US20020037847A1
GENERAL INFORMATION:
APPLICANT: Folkman, M. Judah
APPLICANT: O'Reilly, Michael
TITLE OF INVENTION: Nucleic Acid Encoding Kringle 1-5 Region Fragments of Plasminogen
FILE REFERENCE: 05940-0151 (43171-252088)
CURRENT APPLICATION NUMBER: US/09/761,120
PRIOR APPLICATION NUMBER: 2001-01-16
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 08/866,735
PRIOR FILING DATE: 1997-05-30
NUMBER OF SEQ ID NOS: 47
SOFTWARE: PatentIn version 3.0
SEQ ID NO 30
LENGTH: 250
TYPE: PRT
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Kringle 1-3
US-09-761-120-30

Query Match 11.0%; Score 142; DB 10; Length 250;
Best Local Similarity 37.2%; Pred. No. 2.3e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 4 CFWDNGHLYREDQTSAPAGLRCLNWLDA-----QGLASAPVSGAGNHSYCRNPDEDPFG 58
DB 1 CKTGNGKNYRGTMSTKXGJTCOKMSSTSPHRPSPATHPSEGL-ENYCRNPDPDPOG 59
QY 59 PWCYVSGEAGVPEKR--PCEDLRCEP 82
DB 60 PWCYTTD---PEKRYDYCDILECEE 81

RESULT 15

US-10-131-241-61
Sequence 61 Application US/10131241
Publication No. US20030012792A1
GENERAL INFORMATION:
APPLICANT: Holaday, John W.
APPLICANT: Fortier, Anne H.
TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
FILE REFERENCE: 05213-0344 43170-271565
CURRENT APPLICATION NUMBER: US/10/131,241
PRIOR FILING DATE: 2002-07-22
PRIOR APPLICATION NUMBER: US 09/413,049
PRIOR FILING DATE: 1999-10-06
PRIOR APPLICATION NUMBER: US 09/316,802

PRIOR FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: US 60/086,586
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.1
SEQ ID NO 61
LENGTH: 260
TYPE: PRT
ORGANISM: Homo sapiens
US-10-131-241-61

Query Match 11.0%; Score 142; DB 9; Length 260;
Best Local Similarity 37.2%; Pred. No. 2.4e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 4 CFWDNGHLYREDQTSAPAGLRCLNWLDA-----QGLASAPVSGAGNHSYCRNPDEDPFG 58
DB 6 CKTGNGKNYRGTMSTKXGJTCOKMSSTSPHRPSPATHPSEGL-ENYCRNPDPDPOG 64
QY 59 PWCYVSGEAGVPEKR--PCEDLRCEP 82
DB 65 PWCYTTD---PEKRYDYCDILECEE 86

Search completed: April 7, 2003, 09:31:28
Job time : 18.106 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:17:08 ; Search time 206.782 Seconds

(without alignments)
754,542 Million cell updates/sec

Title: US-10-057-951-2_COPY_22_263

Perfect score: 1289

Sequence: 1 SGGCFMDNGHLHREDQTSRA.....PVDPEGSTRPLMGQAGTREA 242

Scoring table: BLAST62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:*

1: /cgn2_6/ptodata/1/paa/PCTUS.COMB.pep:*

2: /cgn2_6/ptodata/1/paa/US06.COMB.pep:*

3: /cgn2_6/ptodata/1/paa/US07.COMB.pep:*

4: /cgn2_6/ptodata/1/paa/US08.COMB.pep:*

5: /cgn2_6/ptodata/1/paa/US081.COMB.pep:*

6: /cgn2_6/ptodata/1/paa/US082.COMB.pep:*

7: /cgn2_6/ptodata/1/paa/US083.COMB.pep:*

8: /cgn2_6/ptodata/1/paa/US084.COMB.pep:*

9: /cgn2_6/ptodata/1/paa/US085.COMB.pep:*

10: /cgn2_6/ptodata/1/paa/US086.COMB.pep:*

11: /cgn2_6/ptodata/1/paa/US087.COMB.pep:*

12: /cgn2_6/ptodata/1/paa/US088.COMB.pep:*

13: /cgn2_6/ptodata/1/paa/US089.COMB.pep:*

14: /cgn2_6/ptodata/1/paa/US090.COMB.pep:*

15: /cgn2_6/ptodata/1/paa/US091.COMB.pep:*

16: /cgn2_6/ptodata/1/paa/US092.COMB.pep:*

17: /cgn2_6/ptodata/1/paa/US093.COMB.pep:*

18: /cgn2_6/ptodata/1/paa/US094.COMB.pep:*

19: /cgn2_6/ptodata/1/paa/US095.COMB.pep:*

20: /cgn2_6/ptodata/1/paa/US096.COMB.pep:*

21: /cgn2_6/ptodata/1/paa/US097.COMB.pep:*

22: /cgn2_6/ptodata/1/paa/US098.COMB.pep:*

23: /cgn2_6/ptodata/1/paa/US099.COMB.pep:*

24: /cgn2_6/ptodata/1/paa/US100.COMB.pep:*

25: /cgn2_6/ptodata/1/paa/US101.COMB.pep:*

26: /cgn2_6/ptodata/1/paa/US102.COMB.pep:*

27: /cgn2_6/ptodata/1/paa/US60.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | ID | Description |
|------------|-------|--------------------|--------------------------|-------------------|
| 1 | 1289 | 100.0 | 250 14 US-09-001-403-64 | Sequence 64, Appl |
| 2 | 1289 | 100.0 | 263 1 PCT-US00-27239-2 | Sequence 2, Appl |
| 3 | 1289 | 100.0 | 263 1 PCT-US02-05301-164 | Sequence 164, App |
| 4 | 1289 | 100.0 | 263 1 PCT-US98-10728-2 | Sequence 2, Appl |
| 5 | 1289 | 100.0 | 263 14 US-09-084-491A-2 | Sequence 2, Appl |
| 6 | 1289 | 100.0 | 263 15 US-09-144-889-2 | Sequence 2, Appl |

| | | | | |
|----|--------|-------|----------------------------|--------------------|
| 7 | 1289 | 100.0 | 263 24 US-10-057-951-2 | Sequence 2, Appl |
| 8 | 1289 | 100.0 | 263 25 US-10-102-704-2 | Sequence 2, Appl |
| 9 | 1289 | 100.0 | 263 21 US-60-048-000-2 | Sequence 2, Appl |
| 10 | 1286 | 99.8 | 263 27 US-09-791-537-39562 | Sequence 39562, A |
| 11 | 1286 | 99.8 | 263 23 US-09-927-796-44 | Sequence 44, Appl |
| 12 | 1286 | 99.8 | 263 26 US-10-210-951-44 | Sequence 44, Appl |
| 13 | 1286 | 99.8 | 263 26 US-10-211-858-44 | Sequence 44, Appl |
| 14 | 1286 | 99.8 | 263 26 US-10-211-884-44 | Sequence 44, Appl |
| 15 | 1286 | 99.8 | 263 26 US-60-230-435-1580 | Sequence 1580, App |
| 16 | 1286 | 99.8 | 263 27 US-60-389-987-244 | Sequence 244, App |
| 17 | 1286 | 99.8 | 263 27 US-60-412-418-244 | Sequence 244, App |
| 18 | 1275 | 98.9 | 263 20 US-09-611-526-3727 | Sequence 3727, Ap |
| 19 | 1235 | 95.8 | 257 1 PCT-US01-14837-8898 | Sequence 8898, Ap |
| 20 | 1221.5 | 94.8 | 255 1 PCT-US01-14837-8900 | Sequence 8900, Ap |
| 21 | 1166.5 | 90.5 | 308 27 US-60-207-315-471 | Sequence 471, App |
| 22 | 1063.5 | 82.5 | 266 15 US-09-144-889-4 | Sequence 4, Appl |
| 23 | 849 | 65.9 | 178 27 US-60-213-800-237 | Sequence 237, App |
| 24 | 648 | 50.3 | 146 1 PCT-US02-05301-237 | Sequence 237, App |
| 25 | 648 | 50.3 | 146 1 PCT-US02-05301-311 | Sequence 311, App |
| 26 | 596 | 46.2 | 126 17 US-09-307-140-979 | Sequence 979, App |
| 27 | 596 | 46.2 | 126 22 US-09-817-076-979 | Sequence 979, App |
| 28 | 508 | 39.4 | 109 22 US-09-834-366-13839 | Sequence 13839, A |
| 29 | 508 | 39.4 | 109 27 US-60-197-873-13839 | Sequence 13839, A |
| 30 | 476 | 36.9 | 117 1 PCT-US01-14837-8897 | Sequence 8897, Ap |
| 31 | 426 | 33.0 | 88 22 US-09-834-366-16200 | Sequence 16200, A |
| 32 | 426 | 33.0 | 88 27 US-60-197-873-16200 | Sequence 16200, A |
| 33 | 322 | 25.0 | 66 1 PCT-US01-00663-11583 | Sequence 31583, A |
| 34 | 322 | 25.0 | 66 22 US-09-864-761-18457 | Sequence 38457, A |
| 35 | 322 | 25.0 | 66 25 US-10-182-993-30642 | Sequence 30642, A |
| 36 | 322 | 25.0 | 66 25 US-10-182-995-24929 | Sequence 24929, A |
| 37 | 322 | 25.0 | 66 25 US-10-182-997-23626 | Sequence 23626, A |
| 38 | 322 | 25.0 | 66 26 US-10-203-134-13143 | Sequence 31343, A |
| 39 | 322 | 25.0 | 66 26 US-10-203-135-30493 | Sequence 30493, A |
| 40 | 322 | 25.0 | 66 26 US-10-203-136-11400 | Sequence 31400, A |
| 41 | 322 | 25.0 | 66 26 US-10-203-137-31583 | Sequence 31583, A |
| 42 | 322 | 25.0 | 66 26 US-10-203-139-30540 | Sequence 30540, A |
| 43 | 202 | 15.7 | 56 13 US-08-905-051-280 | Sequence 238, App |
| 44 | 202 | 15.7 | 56 19 US-09-547-539C-2338 | Sequence 428, App |
| 45 | 201 | 15.6 | 55 13 US-08-905-135-428 | |

ALIGNMENTS

RESULT 1

US-09-001-403-64

Sequence 64, Application US/09001403

GENERAL INFORMATION:

APPLICANT: Lal, Preeti

APPLICANT: Bardman, Olga

APPLICANT: Hillman, Jennifer L.

APPLICANT: Au-Young, Janice

APPLICANT: Tang, Y. Tom

APPLICANT: Yue, Henry

APPLICANT: Shah, Purvi

APPLICANT: Guegler, Karl J.

APPLICANT: Corley, Neil C.

TITLE OF INVENTION: HUMAN REGULATORY PROTEINS

NUMBER OF SEQUENCES: 150

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US-09/001,403

```

1      FILING DATE:  HERRMETH
2      CLASSIFICATION
3      ATTORNEY/AGENT INFORMATION:
4      NAME:  BILLINGS, LUCY J.
5      REGISTRATION NUMBER:  36,749
6      REFERENCE/DOCKET NUMBER:  PF-045 US
7      TELECOMMUNICATION INFORMATION:
8      TELEPHONE:  (650) 855-0555
9      TELEFAX:  (650) 845-4166
10     INFORMATION FOR SEQ. ID NO. 64:
11     SOURCE CHARACTERISTICS:
12     LENGTH: 250 amino acids
13     TYPE: amino acid
14     STRANDEDNESS: single
15     TOPOLOGY: linear
16     IMMEDIATE SOURCE:
17     LIBRARY: LINDONCOT05
18     CLONE: 312252
19     IS-09-001-403-64

```

| | | | | |
|-----------------------|-----------------|---------------------|-----------|-------------|
| Query Match | 100.0%; | Score 1289; | DB 14; | Length 250; |
| Best Local Similarity | 100.0%; | Pred. No. 1.1e-118; | | |
| Matches 24; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |

| | | | | |
|----|-----|--|---------------------------------------|-----|
| QY | 1 | SGGCFMNDGHLVREDPOTSPA | PELRLCLMWLDQSGLASAPVSGAGNHSXCRLNDEDPG | 60 |
| Db | 9 | SGGCFMNDGHLVREDPOTSPA | PELRLCLMWLDQSGLASAPVSGAGNHSXCRLNDEDPG | 68 |
| QY | 61 | CYVSGAGVPEKRPCEBDLRCEPTTSQALPAFTTEIOEASGSGADPQVFAFANALPARS | | 120 |
| Db | 69 | CYVSGAGVPEKRPCEBDLRCEPTTSQALPAFTTEIOEASGSGADPQVFAFANALPARS | | 128 |
| QY | 121 | EAANAQPIGIGISORVPMMSKEKDLGTGYLVGTTMMVTITAGGILIGSYKRGMDL | | 180 |
| Db | 129 | EAANAQPIGIGISORVPMMSKEKDLGTGYLVGTTMMVTITAGGILIGSYKRGMDL | | 188 |
| QY | 181 | EQHDQVCEKREMRITLPLSAFTNPTCEIVDEKTVVHTSQPVPDQSGSTPLMGQATP | | 240 |
| Db | 189 | EQHDQVCEKREMRITLPLSAFTNPTCEIVDEKTVVHTSQPVPDQSGSTPLMGQATP | | 248 |
| QY | 241 | GA | 242 | |
| Db | 249 | GA | 250 | |

```

RESULT 2
PCT-US00-27239-2
Sequence 2 Application PC/US002027239
GENERAL INFORMATION:
APPLICANT: HUMAN GENOME SCIENCES, INC.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Proteases
FILE REFERENCE: PF318PCT2
CURRENT APPLICATION NUMBER: PCT/US00/27239
CURRENT FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 09/411,977
PRIOR FILING DATE: 1999-10-04
NUMBER OF SEQ. ID NOS: 30
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US00-27239-2

```

| | | | | |
|-----------------------|----------------|---|----------|------------|
| Query Match | 100.0% | Score 1289 | DB 1 | Length 263 |
| Best Local Similarity | 100.0% | Pred. No. 1,2e-118 | | |
| Matches 242 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 |
| Qy | 1 | SGGFGFWNGHLYREDQTSPPAGRCRCNTMTDAOSGLASPVSAGAGNHSYTCRNPDPDPGPM | 60 | |
| | 1 | 1 | | |
| Db | 22 | SGGFGFWNGHLYREDQTSPPAGRCRCNTMTDAOSGLASPVSAGAGNHSYTCRNPDPDPGPM | 81 | |
| Qy | 61 | CYSSEAGVDEKPCEDLRCPETTSQALPAFTTEIQEASBEGPADEVQVAPANALPARS | 120 | |

| | | | | |
|----|--|-----|--|------|
| Dd | | 82 | CYSGEAGPEAKRCPEDLRCPETTSQALPAFTTEIQEASGEAGDAVEQVFAFANNAIPARS | 1414 |
| Qy | | 121 | EAAAVQPVGIGISQRPMNSKKKDLGTLGVLTITMMVIIINAGAILLGSYSRGKDILK | 1809 |
| Dd | | 142 | EAAAVPVGIGISQRPMNSKKKDLGLTGLSVLTITMVIIIINAGIILLGSYSRGDKL | 2010 |
| Qy | | 181 | EQHDKVCEREMQITLPLSAFTNPCEIVDEKTVVVHTSQTPVDPOEGSTPLMGQAATP | 2406 |
| Dd | | 202 | EQHDKVCEREMQITLPLSAFTNPCEIVDEKTVVVHTSQTPVDPOEGSTPLMGQAATP | 2610 |
| Qy | | 241 | GA 242 | |
| Dd | | 262 | GA 263 | |

```

RESULT 3
PCT-US02-05301-164
: Sequence 164, Application PC/TUS0205301
: GENERAL INFORMATION:
: APPLICANT: Human Genome Sciences, Inc.
: TITLE OF INVENTION: 70 Human Secreted Proteins
: FILE REFERENCE: PS736PCT
: CURRENT APPLICATION NUMBER: PCT/US02/05301
: CURRENT FILING DATE: 2002-02-21
: PRIOR APPLICATION NUMBER: US 60/304,417
: PRIOR FILING DATE: 2001-07-12
: PRIOR APPLICATION NUMBER: US 60/270,625
: PRIOR FILING DATE: 2001-02-23
: NUMBER OF SEQ ID NOS: 340
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 164
: LENGTH: 263
: TYPE: PRT
: ORGANISM: Homo sapiens
PCT-US02-05301-164

```

| | | | | | |
|-------------|--------------|--|--------------------|----------|------------|
| Query Match | Similarity | 100.0% | Score 1289 | DE 11 | Length 263 |
| Match Local | Similarity | 100.0% | Pred. No. 1,2e-116 | | |
| Batches 242 | Conservative | 0 | Mismatches 0 | Indels 0 | Gaps 0 |
| QY | 1 | SGCGFMDNGHLVREDOTSPAPGLRCLNMLDQSLASAPVSGAGNHSYCRNDEDPGRGW | 60 | | |
| Db | 22 | SGCGFMDNGHLVREDOTSPAPGLRCLNMLDQSLASAPVSGAGNHSYCRNDEDPGRGW | 81 | | |
| QY | 61 | CYVSGEAGVPEKPCPCDLRCPETTSQALPATPTTIOASGPGADGVQFAPANMLPARS | 120 | | |
| Db | 82 | CYVSGEAGVPEKPCPCDLRCPETTSQALPATPTTIOASGPGADGVQFAPANMLPARS | 141 | | |
| QY | 121 | EAAAVQFVIGISORVAMNSKEKKDGLGVVLGITMMVITIAIGAGIIGSYVKRGKDK | 180 | | |
| Db | 142 | EAAAVQFVIGISORVAMNSKEKKDGLGVVLGITMMVITIAIGAGIIGSYVKRGKDK | 201 | | |
| QY | 181 | EQHDQKVCEREMORITLPASATFNPTCEIYDEKIVVVHVSOTPEVDQEGSTPLMGCAQGP | 240 | | |
| Db | 202 | EQHDQKVCEREMORITLPASATFNPTCEIYDEKIVVVHVSOTPEVDQEGSTPLMGCAQGP | 261 | | |
| QY | 241 | GA 242 | | | |
| Db | 262 | GA 263 | | | |

RESULT 4
PCT-US98-10728-2
Sequence 2, Application PC/TUS9810728
GENERAL INFORMATION:
APPLICANT: EBNER, REINHARD
APPLICANT: MOORE, PAUL
APPLICANT: RUBEN, STEVEN
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.

```

; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/10728
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF379PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8439
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US98-10728-2

Query Match          100.0%; Score 1289; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.2e-118;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFWMDNGLYREDQTSPPAGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGPM 60
DB 22 SGGCFWMDNGLYREDQTSPPAGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGPM 81
QY 61 CVYSGAGVPEKRPCEDLRCPEITTSQALPAFTTEIOEASGPGADEVOVFAPANALPARS 120
DB 82 CVYSGAGVPEKRPCEDLRCPEITTSQALPAFTTEIOEASGPGADEVOVFAPANALPARS 141
QY 121 EAAAVOPVIGISORVRMNSKEKKGDTGLGYVLGITMNVIIIAIGAGIILIGSYKRGKDLK 180
DB 142 EAAAVOPVIGISORVRMNSKEKKGDTGLGYVLGITMNVIIIAIGAGIILIGSYKRGKDLK 201
QY 181 EOHDKVCEREMORITLPLSAFTNPTCEIYDEKTVVHTSQTPVDPQEGSTPLMQAGTP 240
DB 202 EOHDKVCEREMORITLPLSAFTNPTCEIYDEKTVVHTSQTPVDPQEGSTPLMQAGTP 261
QY 241 GA 242
DB 262 GA 263

RESULT 5
US-09-084-491A-2
; Sequence 2, Application US/09084491A
; GENERAL INFORMATION:
; APPLICANT: MOORE, PAUL A.
; APPLICANT: RUBEN, STEVEN M.
; APPLICANT: RUBEN, REINHARD
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/084,491A
; FILING DATE: 27-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF378
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-084-491A-2

Query Match          100.0%; Score 1289; DB 14; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.2e-118;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFWMDNGLYREDQTSPPAGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGPM 60
DB 22 SGGCFWMDNGLYREDQTSPPAGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGPM 81
QY 61 CVYSGAGVPEKRPCEDLRCPEITTSQALPAFTTEIOEASGPGADEVOVFAPANALPARS 120
DB 82 CVYSGAGVPEKRPCEDLRCPEITTSQALPAFTTEIOEASGPGADEVOVFAPANALPARS 141
QY 121 EAAAVOPVIGISORVRMNSKEKKGDTGLGYVLGITMNVIIIAIGAGIILIGSYKRGKDLK 180
DB 142 EAAAVOPVIGISORVRMNSKEKKGDTGLGYVLGITMNVIIIAIGAGIILIGSYKRGKDLK 201
QY 181 EOHDKVCEREMORITLPLSAFTNPTCEIYDEKTVVHTSQTPVDPQEGSTPLMQAGTP 240
DB 202 EOHDKVCEREMORITLPLSAFTNPTCEIYDEKTVVHTSQTPVDPQEGSTPLMQAGTP 261
QY 241 GA 242
DB 262 GA 263

RESULT 6
US-09-144-889-2
; Sequence 2, Application US/09144889B
; GENERAL INFORMATION:
; APPLICANT: EARL F. ALBONE
; APPLICANT: KRISTINE K. KIKLY
; TITLE OF INVENTION: KRINGLE1
; FILE REFERENCE: GH-70249
; CURRENT APPLICATION NUMBER: US/09/144,889B
; CURRENT FILING DATE: 1998-09-01
; EARLIER APPLICATION NUMBER: 60/056,032
; EARLIER FILING DATE: 1997-09-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; TYPE: PRT
; LENGTH: 263
; ORGANISM: HOMO SAPIENS
; US-09-144-889-2

Query Match          100.0%; Score 1289; DB 15; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.2e-118;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFWMDNGLYREDQTSPPAGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGPM 60
DB 22 SGGCFWMDNGLYREDQTSPPAGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGPM 81
```

```
QY 61 CYSGEAGVPEKRPCEDELRCEPTTSOALPAFTTEIOEASGPGADEVOYFAPANALPARS 120
| 82 CYSGEAGVPEKRPCEDELRCEPTTSOALPAFTTEIOEASGPGADEVOYFAPANALPARS 141
DB 121 EAAAVQPVIGISQVRNMSKEKDLGTLGYVLGITMNVIIIAIGAGIILGYSKRGKDLK 180
142 EAAAVQPVIGISQVRNMSKEKDLGTLGYVLGITMNVIIIAIGAGIILGYSKRGKDLK 201
QY 181 EOHDOVCEREMQRTLLPLSAFTNPCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 240
DB 202 EOHDOVCEREMQRTLLPLSAFTNPCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 261
QY 241 GA 242
DB 262 GA 263
```

```
RESULT 7
US-10-057-951-2
; Sequence 2, Application US/10057951
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: PF378P1
; CURRENT APPLICATION NUMBER: US/10/057,951
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 09/411,977
; PRIOR FILING DATE: 1999-10-04
; PRIOR APPLICATION NUMBER: US 09/084,491
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: US 60/048,000
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-951-2
```

```
Query Match 100.0%; Score 1289; DB 24; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.2e-118;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SGGCFWMDNGHLTYREDQTSAPGLRCLNWLDAQSLASAPVSGAGNHSYCRNPDEDPGPGW 60
DB 22 SGGCFWMDNGHLTYREDQTSAPGLRCLNWLDAQSLASAPVSGAGNHSYCRNPDEDPGPGW 81
QY 61 CYSGEAGVPEKRPCEDELRCEPTTSOALPAFTTEIOEASGPGADEVOYFAPANALPARS 120
DB 82 CYSGEAGVPEKRPCEDELRCEPTTSOALPAFTTEIOEASGPGADEVOYFAPANALPARS 141
QY 121 EAAAVQPVIGISQVRNMSKEKDLGTLGYVLGITMNVIIIAIGAGIILGYSKRGKDLK 180
DB 142 EAAAVQPVIGISQVRNMSKEKDLGTLGYVLGITMNVIIIAIGAGIILGYSKRGKDLK 201
QY 181 EOHDOVCEREMQRTLLPLSAFTNPCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 240
DB 202 EOHDOVCEREMQRTLLPLSAFTNPCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 261
QY 241 GA 242
DB 262 GA 263
```

```
RESULT 8
US-10-102-704-2
; Sequence 2, Application US/10102704
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protein
; FILE REFERENCE: PF378C1
```

```
; CURRENT APPLICATION NUMBER: US/10/102,704
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/084,491
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: 60/048,000
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-102-704-2
```

```
Query Match 100.0%; Score 1289; DB 25; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.2e-118;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SGGCFWMDNGHLTYREDQTSAPGLRCLNWLDAQSLASAPVSGAGNHSYCRNPDEDPGPGW 60
DB 22 SGGCFWMDNGHLTYREDQTSAPGLRCLNWLDAQSLASAPVSGAGNHSYCRNPDEDPGPGW 81
QY 61 CYSGEAGVPEKRPCEDELRCEPTTSOALPAFTTEIOEASGPGADEVOYFAPANALPARS 120
DB 82 CYSGEAGVPEKRPCEDELRCEPTTSOALPAFTTEIOEASGPGADEVOYFAPANALPARS 141
QY 121 EAAAVQPVIGISQVRNMSKEKDLGTLGYVLGITMNVIIIAIGAGIILGYSKRGKDLK 180
DB 142 EAAAVQPVIGISQVRNMSKEKDLGTLGYVLGITMNVIIIAIGAGIILGYSKRGKDLK 201
QY 181 EOHDOVCEREMQRTLLPLSAFTNPCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 240
DB 202 EOHDOVCEREMQRTLLPLSAFTNPCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 261
QY 241 GA 242
DB 262 GA 263
```

```
RESULT 9
US-60-048-000-2
; Sequence 2, Application US/60048000
; GENERAL INFORMATION:
; APPLICANT: EBNER, REINHARD
; APPLICANT: MOORE, PAUL
; APPLICANT: RUBEN, STEVE
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSER: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/048,000
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF379PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 2:
```


SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-60-048-000-2

Query Match 100.0%; Score 1289; DB 21; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.2e-118;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFMNGHLVREDQTSAPGLRCINWLDAGSLASAPVSGAGNHSYCRNPEDPRGPW 60
DB 22 SGGCFMNGHLVREDQTSAPGLRCINWLDAGSLASAPVSGAGNHSYCRNPEDPRGPW 81
QY 61 CYSSEAGVPEKRPCEDLRCPTTSQALPAFTTEIOEASRGADAVQVAPANALPARS 120
DB 82 CYSSEAGVPEKRPCEDLRCPTTSQALPAFTTEIOEASRGADAVQVAPANALPARS 141
QY 121 EAAAVQPVIGISQVRMNSKEKKDLGTLGYVIGITMVIILAGIILIGYSYKRGDLK 180
DB 142 EAAAVQPVIGISQVRMNSKEKKDLGTLGYVIGITMVIILAGIILIGYSYKRGDLK 201
QY 181 EHQDQVCEREMORTLPLSAFTNPCEIYDEKTVVHTSQTPVDPEQEGSTPLMGAGTP 240
DB 202 EHQDQVCEREMORTLPLSAFTNPCEIYDEKTVVHTSQTPVDPEQEGSTPLMGAGTP 261
QY 241 GA 242
DB 262 GA 263

RESULT 10

US-09-791-537-39562
Sequence 39562, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Dancier, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patent in version 3.0
SEQ ID NO 39562
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-09-791-537-39562

Query Match 99.8%; Score 1286; DB 21; Length 263;
Best Local Similarity 99.6%; Pred. No. 2.3e-118;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFMNGHLVREDQTSAPGLRCINWLDAGSLASAPVSGAGNHSYCRNPEDPRGPW 60
DB 22 SGGCFMNGHLVREDQTSAPGLRCINWLDAGSLASAPVSGAGNHSYCRNPEDPRGPW 81
QY 61 CYSSEAGVPEKRPCEDLRCPTTSQALPAFTTEIOEASRGADAVQVAPANALPARS 120
DB 82 CYSSEAGVPEKRPCEDLRCPTTSQALPAFTTEIOEASRGADAVQVAPANALPARS 141
QY 121 EAAAVQPVIGISQVRMNSKEKKDLGTLGYVIGITMVIILAGIILIGYSYKRGDLK 180
DB 142 EAAAVQPVIGISQVRMNSKEKKDLGTLGYVIGITMVIILAGIILIGYSYKRGDLK 201
QY 181 EHQDQVCEREMORTLPLSAFTNPCEIYDEKTVVHTSQTPVDPEQEGSTPLMGAGTP 240
DB 202 EHQDQVCEREMORTLPLSAFTNPCEIYDEKTVVHTSQTPVDPEQEGSTPLMGAGTP 261
QY 241 GA 242

DB 262 GA 263

RESULT 11
US-09-927-796-44
Sequence 44, Application US/09927796
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Goddard, Audrey
APPLICANT: Goddard, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Marsden, Scot A.
APPLICANT: Pan, James
APPLICANT: Pitt, Robert M.
APPLICANT: Roy, Margaret Ann
APPLICANT: Smith, Victoria
APPLICANT: Stone, Donna M.
APPLICANT: Watanabe, Colin K.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
FILE REFERENCE: P2931R1C1
CURRENT APPLICATION NUMBER: US/09/927,796
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 60/014699
PRIOR FILING DATE: 1996-04-01
PRIOR APPLICATION NUMBER: 60/026943
PRIOR FILING DATE: 1996-09-23
PRIOR APPLICATION NUMBER: 60/059121
PRIOR FILING DATE: 1997-07-17
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/062037
PRIOR FILING DATE: 1997-10-10
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/067411
PRIOR FILING DATE: 1997-12-03
PRIOR APPLICATION NUMBER: 60/069862
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/095929
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/097978
PRIOR FILING DATE: 1998-08-26
PRIOR APPLICATION NUMBER: 60/103396
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/108867
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/112851
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/119965
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: 60/123972
PRIOR FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: 60/133459
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/140650
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/140653
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20

PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/149395
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/151669
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 08/625328
PRIOR FILING DATE: 1996-04-01
PRIOR APPLICATION NUMBER: 08/710802
PRIOR FILING DATE: 1996-09-23
PRIOR APPLICATION NUMBER: 08/800699
PRIOR FILING DATE: 1997-02-14
PRIOR APPLICATION NUMBER: 08/828683
PRIOR FILING DATE: 1997-03-31
PRIOR APPLICATION NUMBER: 08/829270
PRIOR FILING DATE: 1997-03-31
PRIOR APPLICATION NUMBER: 08/928069
PRIOR FILING DATE: 1997-09-11
PRIOR APPLICATION NUMBER: 08/934494
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 09/143068
PRIOR FILING DATE: 1998-08-28
PRIOR APPLICATION NUMBER: 09/143707
PRIOR FILING DATE: 1998-08-28
PRIOR APPLICATION NUMBER: 09/151889
PRIOR FILING DATE: 1998-09-11
PRIOR APPLICATION NUMBER: 09/169104
PRIOR FILING DATE: 1998-10-09
PRIOR APPLICATION NUMBER: 09/202089
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: 09/254311
PRIOR FILING DATE: 1999-03-03
PRIOR APPLICATION NUMBER: 09/304003
PRIOR FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380139
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423844
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 09/511133
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 09/511631
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 09/664610
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/665350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/690169
PRIOR FILING DATE: 2000-10-16
PRIOR APPLICATION NUMBER: 09/690189
PRIOR FILING DATE: 2000-10-16
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-18
PRIOR APPLICATION NUMBER: 09/866034
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/884733
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: 09/886342
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: 09/866028
PRIOR FILING DATE: 2001-08-25
PRIOR APPLICATION NUMBER: PCT/US97/05230
PRIOR FILING DATE: 1997-03-31
PRIOR APPLICATION NUMBER: PCT/US98/19094

PRIOR FILING DATE: 1998-09-14
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: PCT/US98/21407
PRIOR FILING DATE: 1998-10-09
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: 1998-12-01
PRIOR APPLICATION NUMBER: PCT/US99/05028
PRIOR FILING DATE: 1999-03-08
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/20111
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28301
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/28634
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: PCT/US00/04341
PRIOR FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: PCT/US00/04342
PRIOR FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/06884
PRIOR FILING DATE: 2000-03-15
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: PCT/US00/13705
PRIOR FILING DATE: 2000-05-17
PRIOR APPLICATION NUMBER: PCT/US00/14941
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/20710
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: PCT/US01/17800
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: PCT/US01/19692
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: PCT/US01/21066
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: PCT/US01/21735
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 258
SEQ ID NO 44
LENGTH: 263

Query Match 99.8%; Score 1286; DB 23; Length 263;
Best Local Similarity 99.6%; Pred. No. 2,3e-118;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFWNGHLYREDQTSAPRGRLCLNMLDAQGLASAPVSGAGNHSYCNPDDEDRGPW 60
DB 22 SGGCFWNGHLYREDQTSAPRGRLCLNMLDAQGLASAPVSGAGNHSYCNPDDEDRGPW 81
QY 61 CVYSGEAGVDEKPCEDLRCPETTSOALPAFTTEIOBASGPADEVQVAPANALPARS 120
DB 82 CVYSGEAGVDEKPCEDLRCPETTSOALPAFTTEIOBASGPADEVQVAPANALPARS 141

```
QY 121 EAAVOPVIGISORVRMSKEKDLGTLGYVLGTTMMVITIAIGAGITIGSYKRGKDLK 180
DB 142 EAAVOPVIGISORVRMSKEKDLGTLGYVLGTTMMVITIAIGAGITIGSYKRGKDLK 201
QY 181 EOHDKVCEREMQRTITLPLSAFTNPTCEIVDEKTVVHTSQTVPDPOEGSTPLMGOAGTP 240
DB 202 EOHDKVCEREMQRTITLPLSAFTNPTCEIVDEKTVVHTSQTVPDPOEGSTPLMGOAGTP 261
QY 241 GA 242
DB 262 GA 263
```

RESULT 12

```
US-10-210-951-44
; Sequence 44, Application US/10210951
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Masters, Scott A.
; APPLICANT: Pan, James
; APPLICANT: Picti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/210,951
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-951-44
```

```
Query Match 99.8%; Score 1286; DB 26; Length 263;
Best Local Similarity 99.6%; Pred. No. 2,3e-118;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SGGCFMNDNGHLVRENOTSPAPGLRCLNWLDOQSLASAPVSGAGNHSYCRNPDEDRGFW 60
DB 22 SGGCFMNDNGHLVRENOTSPAPGLRCLNWLDOQSLASAPVSGAGNHSYCRNPDEDRGFW 81
QY 61 CYYSGAGVPEKPCEDLRCPETTSQALPAFTTEIOEASGPGADEVOVFAPANALPARS 120
DB 82 CYYSGAGVPEKPCEDLRCPETTSQALPAFTTEIOEASGPGADEVOVFAPANALPARS 141
```

```
QY 121 EAAVOPVIGISORVRMSKEKDLGTLGYVLGTTMMVITIAIGAGITIGSYKRGKDLK 180
DB 142 EAAVOPVIGISORVRMSKEKDLGTLGYVLGTTMMVITIAIGAGITIGSYKRGKDLK 201
QY 181 EOHDKVCEREMQRTITLPLSAFTNPTCEIVDEKTVVHTSQTVPDPOEGSTPLMGOAGTP 240
DB 202 EOHDKVCEREMQRTITLPLSAFTNPTCEIVDEKTVVHTSQTVPDPOEGSTPLMGOAGTP 261
QY 241 GA 242
DB 262 GA 263
```

RESULT 13

```
US-10-211-858-44
; Sequence 44, Application US/10211858
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Masters, Scott A.
; APPLICANT: Pan, James
; APPLICANT: Picti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/211,858
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-858-44
```

```
Query Match 99.8%; Score 1286; DB 26; Length 263;
Best Local Similarity 99.6%; Pred. No. 2,3e-118;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SGGCFMNDNGHLVRENOTSPAPGLRCLNWLDOQSLASAPVSGAGNHSYCRNPDEDRGFW 60
DB 22 SGGCFMNDNGHLVRENOTSPAPGLRCLNWLDOQSLASAPVSGAGNHSYCRNPDEDRGFW 81
QY 61 CYYSGAGVPEKPCEDLRCPETTSQALPAFTTEIOEASGPGADEVOVFAPANALPARS 120
```

Db 82 CYSVSGAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASGEGADEVQYFAPANALPARS 141
QY 121 EAAAVQPVIGISQVRNMSKEKKDLGTLGYVLGITMMVIIAIGAGIILGYSYKRGKDLK 180
Db 142 EAAAVQPVIGISQVRNMSKEKKDLGTLGYVLGITMMVIIAIGAGIILGYSYKRGKDLK 201
QY 181 EOHDKVCEREMQRTITPLSAFTNPCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 240
Db 202 EOHDKVCEREMQRTITPLSAFTNPCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 261
QY 241 GA 242
Db 262 GA 263

RESULT 14
US-10-211-884-44
; Sequence 44, Application US/10211884
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Pizzi, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Matanabe, Colin K.
; APPLICANT: Wood, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/211,884
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-884-44

Query Match 99.8%; Score 1286; DB 26; Length 263;
Best Local Similarity 99.6%; Pred. No. 2,3e-118;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 14 SGCCFMDNGHLTYREDQTSAPGLRCLNWLDAQSLASAPVSGAGNHSYCRNPDEDPRGPW 60
Db 22 SGCCFMDNGHLTYREDQTSAPGLRCLNWLDAQSLASAPVSGAGNHSYCRNPDEDPRGPW 81
QY 61 CYSVSGAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASGEGADEVQYFAPANALPARS 120

Db 82 CYSVSGAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASGEGADEVQYFAPANALPARS 141
QY 121 EAAAVQPVIGISQVRNMSKEKKDLGTLGYVLGITMMVIIAIGAGIILGYSYKRGKDLK 180
Db 142 EAAAVQPVIGISQVRNMSKEKKDLGTLGYVLGITMMVIIAIGAGIILGYSYKRGKDLK 201
QY 181 EOHDKVCEREMQRTITPLSAFTNPCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 240
Db 202 EOHDKVCEREMQRTITPLSAFTNPCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 261
QY 241 GA 242
Db 262 GA 263

RESULT 15
US-60-230-435-1580
; Sequence 1580, Application US/60230435
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: C1000768
; CURRENT APPLICATION NUMBER: US/60/230,435
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 2991
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1580
; LENGTH: 263
; TYPE: PRT
; ORGANISM: HUMAN
US-60-230-435-1580

Query Match 99.8%; Score 1286; DB 27; Length 263;
Best Local Similarity 99.6%; Pred. No. 2,3e-118;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGCCFMDNGHLTYREDQTSAPGLRCLNWLDAQSLASAPVSGAGNHSYCRNPDEDPRGPW 60
Db 22 SGCCFMDNGHLTYREDQTSAPGLRCLNWLDAQSLASAPVSGAGNHSYCRNPDEDPRGPW 81
QY 61 CYSVSGAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASGEGADEVQYFAPANALPARS 120
Db 82 CYSVSGAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASGEGADEVQYFAPANALPARS 141
QY 121 EAAAVQPVIGISQVRNMSKEKKDLGTLGYVLGITMMVIIAIGAGIILGYSYKRGKDLK 180
Db 142 EAAAVQPVIGISQVRNMSKEKKDLGTLGYVLGITMMVIIAIGAGIILGYSYKRGKDLK 201
QY 181 EOHDKVCEREMQRTITPLSAFTNPCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 240
Db 202 EOHDKVCEREMQRTITPLSAFTNPCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 261
QY 241 GA 242
Db 262 GA 263

Search completed: April 7, 2003, 09:29:26
Job time : 207.782 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Comogen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:18:38 ; Search time 36.2245 Seconds
(without alignments)
1010.317 Million cell updates/sec

Title: US-10-057-951-2_COPY_22_263

Perfect score: 1289
Sequence: 1 SGCGCFMDNGHLYREDQTSAPA.....PVDQEGSTPLMGQATPGA 242

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 679521 seqs, 151232488 residues

Total number of hits satisfying chosen parameters: 679521

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*

1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pep:*\n2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep:*\n3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep:*\n4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep:*\n5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:*\n6: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep:*\n7: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---------------------|
| 1 | 1280 | 99.3 | 263 | 6 | US-10-218-140-6002 |
| 2 | 1028.5 | 79.8 | 264 | 6 | US-10-144-779-488 |
| 3 | 154.5 | 12.0 | 293 | 7 | US-60-452-680-14406 |
| 4 | 154.5 | 12.0 | 293 | 7 | US-60-453-135-8958 |
| 5 | 154.5 | 12.0 | 293 | 7 | US-60-453-050-8958 |
| 6 | 154.5 | 12.0 | 516 | 7 | US-60-452-680-14405 |
| 7 | 154.5 | 12.0 | 516 | 7 | US-60-453-135-8957 |
| 8 | 154.5 | 12.0 | 516 | 7 | US-60-453-050-8957 |
| 9 | 154.5 | 12.0 | 527 | 5 | US-09-612-314A-51 |
| 10 | 154.5 | 12.0 | 562 | 1 | PCT-US02-32263-26 |
| 11 | 154.5 | 12.0 | 562 | 5 | US-09-703-695A-4 |
| 12 | 154.5 | 12.0 | 562 | 6 | US-10-287-994-26 |
| 13 | 154.5 | 12.0 | 562 | 7 | US-60-452-680-14408 |
| 14 | 154.5 | 12.0 | 562 | 7 | US-60-453-135-8960 |
| 15 | 154.5 | 12.0 | 562 | 7 | US-60-453-050-8960 |
| 16 | 154.5 | 12.0 | 587 | 5 | US-09-949-016-11501 |
| 17 | 146.5 | 11.4 | 655 | 1 | PCT-US02-19017-28 |
| 18 | 146.5 | 11.4 | 655 | 7 | US-60-453-135-11516 |
| 19 | 146.5 | 11.4 | 656 | 7 | US-60-453-050-11516 |
| 20 | 146 | 11.3 | 322 | 1 | PCT-US02-27855-21 |
| 21 | 146 | 11.3 | 322 | 1 | PCT-US02-27855A-21 |
| 22 | 143.5 | 11.1 | 322 | 6 | US-10-233-675A-21 |
| 23 | 143.5 | 11.1 | 482 | 7 | US-60-452-680-14407 |
| 24 | 143.5 | 11.1 | 482 | 7 | US-60-453-135-8959 |
| 25 | 143.5 | 11.1 | 482 | 7 | US-60-453-050-8959 |
| 26 | 142 | 11.0 | 339 | 6 | US-10-401-108-3 |

| | | | | | | |
|----|-------|------|-----|---|---------------------|-------------------|
| 27 | 142 | 11.0 | 363 | 6 | US-10-292-418-11 | Sequence 11, Appl |
| 28 | 142 | 11.0 | 391 | 1 | PCT-US02-37879-7 | Sequence 7, Appl |
| 29 | 142 | 11.0 | 391 | 6 | US-10-304-287-7 | Sequence 7, Appl |
| 30 | 142 | 11.0 | 394 | 1 | PCT-US02-37879-8 | Sequence 8, Appl |
| 31 | 142 | 11.0 | 394 | 6 | US-10-304-287-8 | Sequence 8, Appl |
| 32 | 142 | 11.0 | 453 | 5 | US-09-291-200A-1 | Sequence 1, Appl |
| 33 | 142 | 11.0 | 453 | 5 | US-09-946-893B-4 | Sequence 4, Appl |
| 34 | 142 | 11.0 | 569 | 5 | US-09-946-893B-5 | Sequence 5, Appl |
| 35 | 142 | 11.0 | 571 | 5 | US-09-946-893B-8 | Sequence 8, Appl |
| 36 | 142 | 11.0 | 576 | 5 | US-09-946-893B-6 | Sequence 6, Appl |
| 37 | 142 | 11.0 | 713 | 5 | US-09-949-016-9983 | Sequence 9983, Ap |
| 38 | 142 | 11.0 | 713 | 7 | US-60-452-680-22399 | Sequence 22399, A |
| 39 | 142 | 11.0 | 713 | 7 | US-60-453-135-13847 | Sequence 13847, A |
| 40 | 142 | 11.0 | 713 | 7 | US-60-453-050-13847 | Sequence 13847, A |
| 41 | 142 | 11.0 | 791 | 1 | PCT-US02-37879-1 | Sequence 1, Appl |
| 42 | 142 | 11.0 | 791 | 6 | US-10-304-287-1 | Sequence 1, Appl |
| 43 | 142 | 11.0 | 810 | 5 | US-09-291-200A-5 | Sequence 5, Appl |
| 44 | 142 | 11.0 | 810 | 5 | US-09-946-893B-2 | Sequence 2, Appl |
| 45 | 135.5 | 10.5 | 688 | 1 | PCT-US02-27855-18 | Sequence 18, Appl |

ALIGNMENTS

RESULT 1
US-10-218-140-6002
Sequence 6002, Application US/10218140
GENERAL INFORMATION:
APPLICANT: Leach, Martin D.
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES
FILE REFERENCE: 15966-543 CON
CURRENT FILING DATE: 2002-08-12
PRIOR APPLICATION NUMBER: 09/540,763
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: 60/127,728
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/127,636
PRIOR FILING DATE: 1999-04-02
PRIOR APPLICATION NUMBER: 60/127,607
PRIOR FILING DATE: 1999-03-31
NUMBER OF SEQ ID NOS: 6322
SOFTWARE: Curator Version 1.0
SEQ ID NO 6002
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-10-218-140-6002

Query Match 99.3%; Score 1280; DB 6; Length 263;
Best Local Similarity 99.2%; Pred. No. 5.4e-102;
Matches 240; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 1 | SGCGCFMDNGHLYREDQTSAPAGLCLMWLDQGLASAPVSGAGNHSYCRNDPEPRGFW | 60 |
| DB | 22 | SGCGCFMDNGHLYREDQTSAPAGLCLMWLDQGLASAPVSGAGNHSYCRNDPEPRGFW | 81 |
| QY | 61 | CVSGEAGVPEKRPEDIRCPETTSOALPAFTTEIOEASGEPADBEVOVFAPANALPARS | 120 |
| DB | 82 | CVSGEAGVPEKRPEDIRCPETTSOALPAFTTEIOEASGEPADBEVOVFAPANALPARS | 141 |
| QY | 121 | EAAAVQPIYIGISQVRNRSKKKDLGTLGYLVGLTMMVITIAIGAGIILIGYSYKKGKDKL | 180 |
| DB | 142 | EAAAVQPIYIGISQVRNRSKKKDLGTLGYLVGLTMMVITIAIGAGIILIGYSYKKGKDKL | 201 |
| QY | 181 | EOHDQKVEREMQRTTLPISAFNPTCEIYDEKTVVHTSQTFVVDQEGSTPLMGQATGP | 240 |
| DB | 202 | EOHDQKVEREMQRTTLPISAFNPTCEIYDEKTVVHTSQTFVVDQEGSTPLMGQATGP | 261 |
| QY | 241 | GA 242 | |

Db 262 GA 263

RESULT 2

US-10-144-779-488
; Sequence 488, Application US/10144779
; GENERAL INFORMATION:
; APPLICANT: SUBRAMANIAN, Mani et al.
; TITLE OF INVENTION: MOUSE ORTHOLOGS OF HUMAN DISEASE GENES.
; FILE REFERENCE: CL001235
; CURRENT APPLICATION NUMBER: US/10/144,779
; CURRENT FILING DATE: 2002-05-15
; NUMBER OF SEQ ID NOS: 864
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 488
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-144-779-488

Query Match 79.8%; Score 1028.5; DB 6; Length 264;
Best Local Similarity 78.6%; Pred. No. 2.2e-80;
Matches 191; Conservative 16; Mismatches 35; Indels 1; Gaps 1;

QY 1 SGGCFDNGHLYREDQTSAPGRLCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPKPGW 60
DB 22 SGGCFDNGHLYREDQTSAPGRLCLNWLDAQSGRESLTPSPGNHNYCRNPDDPGRPW 81
QY 61 CYSVGENGVPEKRPCEDLRCPEPTTSQA-LPATTETIQEASGEGPADDEVQVAPANALPAR 119
DB 82 CYSVSEGTGVEKRPCEDLRCPEPTTSQAPPPSSAMELEKSGAPDKAQPVPANALPAR 141
QY 120 SEAAVQVPVIGISQRYRMSKEKDLGTGLGYLGITTMVYIIAGGIIIGYSYRGKDL 179
DB 142 SEAAVQVPVIGISQRYRMSKEKDLGTGLGYLGITTMVYIIAGGIIIGYSYRGKDL 201
QY 180 KECHDKVCEREMORITLPLSAFTNFTCEIIDEKTVVHTSQTPVDPEGSTPLMGAGT 239
DB 202 KECHDKVCEREMORITLPLSAFTNFTCEIIDEKTVVHTSQTPVDPEGSTPLMGAGT 261
QY 240 PGA 242
DB 262 PGA 264

RESULT 3

US-60-452-680-14406
; Sequence 14406, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GROPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14406
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-452-680-14406

Query Match 12.0%; Score 154.5; DB 7; Length 293;
Best Local Similarity 39.6%; Pred. No. 2.8e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFWDNGHLYREDQTSAPGRLCLNWLDAQSGLASAPVSGAGNHSYCRNPDE 54
DB 127 CYEDGISTYRGWSTAESGAECTNM--NSSALAQKPYSGRRPDALIRGLGNHNYCRNPDR 184

QY 55 DPGFPCYVSGEAGVPEKRPCEDLRCPEPTTS 85
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213

RESULT 4

US-60-453-135-8958
; Sequence 8958, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8958
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-8958

Query Match 12.0%; Score 154.5; DB 7; Length 293;
Best Local Similarity 39.6%; Pred. No. 2.8e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFWDNGHLYREDQTSAPGRLCLNWLDAQSGLASAPVSGAGNHSYCRNPDE 54
DB 127 CYEDGISTYRGWSTAESGAECTNM--NSSALAQKPYSGRRPDALIRGLGNHNYCRNPDR 184
QY 55 DPGFPCYVSGEAGVPEKRPCEDLRCPEPTTS 85
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213

RESULT 5

US-60-453-050-8958
; Sequence 8958, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8958
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-8958

Query Match 12.0%; Score 154.5; DB 7; Length 293;
Best Local Similarity 39.6%; Pred. No. 2.8e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFWDNGHLYREDQTSAPGRLCLNWLDAQSGLASAPVSGAGNHSYCRNPDE 54
DB 127 CYEDGISTYRGWSTAESGAECTNM--NSSALAQKPYSGRRPDALIRGLGNHNYCRNPDR 184
QY 55 DPGFPCYVSGEAGVPEKRPCEDLRCPEPTTS 85
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213

RESULT 6

US-60-452-680-14405
; Sequence 14405, Application US/60452680
; GENERAL INFORMATION:

```

; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CLO01450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 14405
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-452-680-14405

Query Match
Best Local Similarity 12.0%; Score 154.5; DB 7; Length 516;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFWDNGHLHYREDOTSPAPGLRCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 54
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 81 CYEDOGISYRGWTWSTAESGAECTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCRNPDR 138
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 55 DPRGPMCYVSGEAGVPEKRPCEDLRCPEPTTS 85
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 139 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 167

RESULT 7
US-60-453-135-8957
; Sequence 8957, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CLO01456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8957
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-8957

Query Match
Best Local Similarity 12.0%; Score 154.5; DB 7; Length 516;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFWDNGHLHYREDOTSPAPGLRCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 54
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 81 CYEDOGISYRGWTWSTAESGAECTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCRNPDR 138
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 55 DPRGPMCYVSGEAGVPEKRPCEDLRCPEPTTS 85
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 139 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 167

RESULT 8
US-60-453-050-8957
; Sequence 8957, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CLO01457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
```

```

; SEQ ID NO 8957
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-8957

Query Match
Best Local Similarity 12.0%; Score 154.5; DB 7; Length 516;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFWDNGHLHYREDOTSPAPGLRCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 54
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 81 CYEDOGISYRGWTWSTAESGAECTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCRNPDR 138
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 55 DPRGPMCYVSGEAGVPEKRPCEDLRCPEPTTS 85
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 139 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 167

RESULT 9
US-09-612-314A-51
; Sequence 51, Application US/09612314A
; GENERAL INFORMATION:
; APPLICANT: SMITH, RICHARD ANTHONY GODWIN
; APPLICANT: DODD, IAN
; APPLICANT: MOSAKOWSKA, DANUTA EWA IRENA
; TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
; FILE REFERENCE: 37945-0004
; CURRENT APPLICATION NUMBER: US/09/612,314A
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 09/214,913
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: PCT/EP97/03715
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: GB 96 148 71.3
; PRIOR FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: tissue plasminogen activator
US-09-612-314A-51

Query Match
Best Local Similarity 12.0%; Score 154.5; DB 5; Length 527;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFWDNGHLHYREDOTSPAPGLRCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 54
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 92 CYEDOGISYRGWTWSTAESGAECTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCRNPDR 149
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 55 DPRGPMCYVSGEAGVPEKRPCEDLRCPEPTTS 85
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 178
```

```

RESULT 10
PCT-US02-32263-26
; Sequence 26, Application PC/TUS0232263
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DePrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Bove, Garyn
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; TITLE OF INVENTION: REMODELING AND GLYCOCONJUGATION OF PEPTIDES
; FILE REFERENCE: 040853-01-5050WO
```


Db 127 CYEDGISTYRGWSTAESGAECTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCRNPDR 184

QY 55 DRGPMWCYVSGEAGVPEKRPCEDLRCPEPTTS 85
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213

RESULT 14
US-60-453-135-8960

; Sequence 8960, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8960
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-8960

Query Match 12.0%; Score 154.5; DB 7; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.1e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFWNDGHLVREDQTSPPAPLRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 54
Db 127 CYEDGISTYRGWSTAESGAECTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCRNPDR 184

QY 55 DRGPMWCYVSGEAGVPEKRPCEDLRCPEPTTS 85
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213

RESULT 15
US-60-453-050-8960

; Sequence 8960, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8960
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-8960

Query Match 12.0%; Score 154.5; DB 7; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.1e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFWNDGHLVREDQTSPPAPLRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 54
Db 127 CYEDGISTYRGWSTAESGAECTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCRNPDR 184

QY 55 DRGPMWCYVSGEAGVPEKRPCEDLRCPEPTTS 85
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213

Search completed: April 7, 2003, 09:30:47

Job time : 37.2245 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:28 ; Search time 24.1497 Seconds
(without alignments)
294.842 Million cell updates/sec

Title: US-10-057-951-2_COPY_22_263
Perfect score: 1289
Sequence: 1 SGGCFMNGHLYREDTSP...PVDPOEGSTPLMGAGTGA 242

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_Aa:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1 | 1289 | 100.0 | 263 | 4 US-09-411-977-2 | Sequence 2, Appli |
| 2 | 158.5 | 12.3 | 472 | 2 US-08-811-949-63 | Sequence 63, Appl |
| 3 | 154.5 | 12.0 | 437 | 2 US-08-811-949-49 | Sequence 49, Appl |
| 4 | 154.5 | 12.0 | 437 | 2 US-08-811-949-51 | Sequence 51, Appl |
| 5 | 154.5 | 12.0 | 437 | 2 US-08-811-949-55 | Sequence 55, Appl |
| 6 | 154.5 | 12.0 | 437 | 2 US-08-811-949-57 | Sequence 57, Appl |
| 7 | 154.5 | 12.0 | 527 | 1 US-07-609-510B-16 | Sequence 16, Appl |
| 8 | 154.5 | 12.0 | 527 | 2 US-08-811-949-39 | Sequence 39, Appl |
| 9 | 154.5 | 12.0 | 527 | 5 PCR-US91-01025A-2 | Sequence 2, Appli |
| 10 | 154.5 | 12.0 | 527 | 6 5185259-8 | Patent No. 5185259 |
| 11 | 154.5 | 12.0 | 527 | 6 5520913-1 | Patent No. 5520913 |
| 12 | 154.5 | 12.0 | 546 | 6 5200340-6 | Patent No. 5200340 |
| 13 | 154.5 | 12.0 | 562 | 2 US-08-811-949-43 | Sequence 43, Appl |
| 14 | 154.5 | 12.0 | 562 | 2 US-08-560-098A-50 | Sequence 50, Appl |
| 15 | 154.5 | 12.0 | 562 | 2 US-08-883-795A-38 | Sequence 38, Appl |
| 16 | 154.5 | 12.0 | 562 | 6 5185259-3 | Patent No. 5185259 |
| 17 | 154.5 | 12.0 | 562 | 6 5200340-2 | Patent No. 5200340 |
| 18 | 154.5 | 12.0 | 562 | 6 5344773-2 | Patent No. 5344773 |
| 19 | 148.5 | 11.5 | 83 | 2 US-08-811-949-2 | Sequence 2, Appli |
| 20 | 146.5 | 11.4 | 655 | 2 US-08-148-910-12 | Sequence 12, Appl |
| 21 | 146.5 | 11.4 | 655 | 1 US-08-448-937A-12 | Sequence 12, Appl |
| 22 | 142 | 11.0 | 160 | 2 US-08-612-788-35 | Sequence 35, Appl |
| 23 | 142 | 11.0 | 160 | 2 US-09-066-028-35 | Sequence 35, Appl |
| 24 | 142 | 11.0 | 250 | 2 US-08-612-788-30 | Sequence 30, Appl |
| 25 | 142 | 11.0 | 250 | 3 US-09-066-028-30 | Sequence 30, Appl |
| 26 | 142 | 11.0 | 339 | 1 US-08-248-629A-3 | Sequence 3, Appli |
| 27 | 142 | 11.0 | 339 | 1 US-08-451-932-3 | Sequence 3, Appli |

| | | | | | |
|----|-----|------|-----|--------------------|--------------------|
| 28 | 142 | 11.0 | 339 | 1 US-08-452-260-3 | Sequence 3, Appli |
| 29 | 142 | 11.0 | 339 | 1 US-08-326-785-3 | Sequence 3, Appli |
| 30 | 142 | 11.0 | 339 | 2 US-08-612-788-3 | Sequence 3, Appli |
| 31 | 142 | 11.0 | 339 | 2 US-08-605-598B-3 | Sequence 3, Appli |
| 32 | 142 | 11.0 | 339 | 2 US-08-429-743-3 | Sequence 3, Appli |
| 33 | 142 | 11.0 | 339 | 2 US-08-866-735-3 | Sequence 3, Appli |
| 34 | 142 | 11.0 | 339 | 3 US-09-066-028-3 | Sequence 3, Appli |
| 35 | 142 | 11.0 | 339 | 5 PCR-US95-05107-3 | Sequence 3, Appli |
| 36 | 142 | 11.0 | 352 | 2 US-08-612-788-40 | Sequence 40, Appl |
| 37 | 142 | 11.0 | 352 | 3 US-09-066-028-40 | Sequence 40, Appl |
| 38 | 142 | 11.0 | 374 | 4 US-09-377-250-3 | Sequence 3, Appli |
| 39 | 142 | 11.0 | 375 | 4 US-09-377-250-2 | Sequence 2, Appli |
| 40 | 142 | 11.0 | 378 | 2 US-08-612-788-42 | Sequence 42, Appl |
| 41 | 142 | 11.0 | 378 | 4 US-09-066-028-42 | Sequence 42, Appl |
| 42 | 142 | 11.0 | 378 | 4 US-09-206-059-1 | Sequence 1, Appli |
| 43 | 142 | 11.0 | 451 | 4 US-09-377-250-1 | Sequence 1, Appli |
| 44 | 142 | 11.0 | 452 | 4 US-09-377-250-4 | Sequence 4, Appli |
| 45 | 142 | 11.0 | 790 | 1 US-08-469-486-54 | Sequence 54, Appli |

ALIGNMENTS

RESULT 1
US-09-411-977-2
Sequence 2, Application US/09411977
Patent No. 6372473
GENERAL INFORMATION:
APPLICANT: Moore, Paul A.
APPLICANT: Kuber, Steven M.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
FILE REFERENCE: PF378P1
CURRENT APPLICATION NUMBER: US/09/411,977
CURRENT FILING DATE: 1999-10-04
EARLIER APPLICATION NUMBER: 09/084,491
EARLIER FILING DATE: 1998-05-27
EARLIER APPLICATION NUMBER: 60/048,000
EARLIER FILING DATE: 1997-05-28
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-09-411-977-2

Query Match 100.0%; Score 1289; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 9.9e-129;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 1 | SGGCFMNGHLYREDTSPAPGLRCINWLDASGLASAPVSGAGNHSYCRNPDEDRGFW | 60 |
| DB | 22 | SGGCFMNGHLYREDTSPAPGLRCINWLDASGLASAPVSGAGNHSYCRNPDEDRGFW | 81 |
| QY | 61 | CYVSGEAGVEKPCEDLRCPETTSQALPAFTTEIOEASGEGADEVQVAPANALPARS | 120 |
| DB | 82 | CYVSGEAGVEKPCEDLRCPETTSQALPAFTTEIOEASGEGADEVQVAPANALPARS | 141 |
| QY | 121 | EAAAVQPVISQVRVMSKEKDLGLGYVIGITMVIITAGIITIGSYKRRKDK | 180 |
| DB | 142 | EAAAVQPVISQVRVMSKEKDLGLGYVIGITMVIITAGIITIGSYKRRKDK | 201 |
| QY | 181 | BOHDQKCEKEMRITLPLSAFTNPCEIYDEKTVVHNSQTPVDEOEGSTPLMGAGTGA | 240 |
| DB | 202 | BOHDQKCEKEMRITLPLSAFTNPCEIYDEKTVVHNSQTPVDEOEGSTPLMGAGTGA | 261 |
| QY | 241 | GA 242 | |
| DB | 262 | GA 263 | |

RESULT 2

OY 4 CFWNDGHIYREDQTSAPAPGIRCLTNLDKQSGLASAVS-----GAQNHSYCNDE 54
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 92 CYEDDGISIYGRTWSTASGEECTNW--NSSLALNQKYRSRRPDAIRJGLGNHHYCNDNR 148
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 55 DPGRMVYSGSEAGVPEPKPCPDLCACPETHS 85
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 150 DSK-PMCTYF-KAGKYSSEFCSTPAICSGNS 176
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 10
5185259-8
Patent No. 5185259
APPLICANT: GORDELL, DAVID V., KOHR, WILLIAM J., PENNICA, DIANE,
VEHAR, GORDON A.
TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
ACTIVATOR
NUMBER OF SEQUENCES: 15
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/489,855
FILING DATE: 02-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 12,694
FILING DATE: 09-FEB-1987
APPLICATION NUMBER: 483,052
FILING DATE: 07-APR-1982
APPLICATION NUMBER: 398,003
FILING DATE: 14-JUL-1982
APPLICATION NUMBER: 374,860
FILING DATE: 05-MAY-1982
SEQ ID NO: 8
LENGTH: 527
5185259-8

```

RESULT 11
5520913-1
; Patent No. 5520913
; APPLICANT: ANDERSON, STEPHEN; BENNETT, WILLIAM F.; BOTSSTEIN,
; DAVID; HIGGINS, DEBORAH L.; PAONI, NICHOLAS F.; ZOLLER, MARK J.
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR HAVING
; ZYMOCINIC PROPERTIES
; NUMBER OF SEQUENCES: 35
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/88,451
; FILING DATE: 06 JUL-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 770,510
; FILING DATE: 03-OCT-1991
; APPLICATION NUMBER: 384,608
; FILING DATE: 24-JUL-1989
; APPLICATION NUMBER: 240,856
; FILING DATE: 02-SEP-1988
; SEQ ID NO:1;
; LENGTH:527
; 5520913-1

```

```

Oy  4  CFWDDGHHLYEEDQTSAPRLRLCLYMLDAOSGLASAPVS-----GAGNHSYRANDE  54
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db  92  CYEDGGISYRGTSTWSTMSGAECTNN--NSALALQKYSRRPDAIRLGGGNNHYCRNDR  149
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Oy  55  DPGRCWCVSGEAGVGRBKPRCDLRCPETTS  85
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db  150  DSK-PACTVF-KAGKTSSEFCSTPACSGNS  178

```

```

RESULT 12
5200340-6
: Patent No. 5200340
: APPLICANT: FOSTER, DONALD C.,MULVIHILL, EILEEN R.,O'HARA,
: PATRICK J.,PENGEL, KIRTY,YOSHITAKE, SHINJI
: TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
: ACTIVATORS
: NUMBER OF SEQUENCES: 34
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/53,412
: FILING DATE: 22-MAY-1987
: SEQ ID NO:6:
: LENGTH: 546
5200340-6

```

US-RESULT 13
US-08-911-949-43
Sequence 43, Application US/0681949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIMA, MINO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASARAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBOLON, SPIVAK, MCCLELLAND, MAIER &
ADDRESSER: P. C.
STEERY: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBOLON, NORMAN P.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-43

Query Match 12.0%; Score 154.5; DB 2; Length 562;
Best Local Similarity 39.6%; Pred. No. 7.3e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFWDNGHYREDQTSAPGLRCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 54
DB 127 CYEDGISTYRGWTSTAGSAGACTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCRNPDR 184
QY 55 DPRGWCYSGEAGVPEKRPCEDLACPETS 85
DB 185 DSK-FWCYVF-KAGKYSSEFCSTPACSEGN 213

RESULT 14
US-08-560-098A-50
Sequence 50, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WENDET, Stephan
APPLICANT: HEINZEL, WILFRIED, Regina
TITLE OF INVENTION: Proteins having fibrinolytic and
NUMBER OF INVENTION: Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-50

Query Match 12.0%; Score 154.5; DB 2; Length 562;
Best Local Similarity 39.6%; Pred. No. 7.3e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
QY 4 CFWDNGHYREDQTSAPGLRCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 54
DB 127 CYEDGISTYRGWTSTAGSAGACTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCRNPDR 184

DB 127 CYEDGISTYRGWTSTAGSAGACTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCRNPDR 184
QY 55 DPRGWCYSGEAGVPEKRPCEDLACPETS 85
DB 185 DSK-FWCYVF-KAGKYSSEFCSTPACSEGN 213

RESULT 15
US-08-883-795A-38
Sequence 38, Application US/08883795A
Patent No. 5985607
GENERAL INFORMATION:
APPLICANT: Delcive, Genevieve
APPLICANT: Awang, Gregor
TITLE OF INVENTION: Recombinant DNA Molecules and Expression
NUMBER OF INVENTION: Vectors for Tissue Plasminogen Activator
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,795A
FILING DATE: 27-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Micheline
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 7841-062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein (tpa)
US-08-883-795A-38

Query Match 12.0%; Score 154.5; DB 2; Length 562;
Best Local Similarity 39.6%; Pred. No. 7.3e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
QY 4 CFWDNGHYREDQTSAPGLRCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 54
DB 127 CYEDGISTYRGWTSTAGSAGACTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCRNPDR 184
QY 55 DPRGWCYSGEAGVPEKRPCEDLACPETS 85
DB 185 DSK-FWCYVF-KAGKYSSEFCSTPACSEGN 213

Search completed: April 7, 2003, 09:22:29
Job time : 26.1497 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:25 ; Search time 4.86486 Seconds

(without alignments)
1185.658 Million cell updates/sec

Title: US-10-057-951-2_COPY_25_84

Perfect score: 357

Sequence: 1 CFMDNGHLYREDQTSAPAGL.....GNHSYCRNPEDRGPWCYV 60

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.73.*
2: PIR1.*
3: PIR2.*
4: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | ID | Description |
|------------|-------|--------------------|---------------|--------------------|
| 1 | 146 | 40.9 | 562 1 UKHUT | t-plasminogen acti |
| 2 | 141 | 39.5 | 559 1 A35029 | t-plasminogen acti |
| 3 | 141 | 39.5 | 655 1 A46688 | hepatocyte growth |
| 4 | 137 | 38.4 | 291 2 I38098 | t-plasminogen acti |
| 5 | 136 | 38.1 | 559 1 A29941 | t-plasminogen acti |
| 6 | 134 | 37.5 | 431 2 JS0599 | t-plasminogen acti |
| 7 | 134 | 37.5 | 477 1 A34369 | t-plasminogen acti |
| 8 | 134 | 37.5 | 477 2 JS0598 | t-plasminogen acti |
| 9 | 132 | 37.0 | 433 1 UKBAY | u-plasminogen acti |
| 10 | 131 | 36.7 | 431 1 UKHU | u-plasminogen acti |
| 11 | 130 | 36.4 | 442 1 UKPG | u-plasminogen acti |
| 12 | 128 | 35.9 | 716 1 JCS061 | u-plasminogen acti |
| 13 | 127 | 35.6 | 433 1 JN0560 | macrophage-stimula |
| 14 | 126 | 35.3 | 394 2 JS0600 | t-plasminogen acti |
| 15 | 126 | 35.3 | 716 1 A40332 | macrophage-stimula |
| 16 | 123 | 34.5 | 169 2 A40522 | plasmin (EC 3.4.21 |
| 17 | 121 | 33.9 | 810 1 PLHU | plasmin (EC 3.4.21 |
| 18 | 119 | 33.3 | 433 1 UKMS | u-plasminogen acti |
| 19 | 119 | 33.3 | 728 1 JH0579 | hepatocyte growth |
| 20 | 118 | 33.1 | 432 1 SI8932 | u-plasminogen acti |
| 21 | 118 | 33.1 | 728 1 A60185 | hepatocyte growth |
| 22 | 118 | 33.1 | 810 2 I46260 | plasmin (EC 3.4.21 |
| 23 | 115 | 32.2 | 477 2 JS0597 | t-plasminogen acti |
| 24 | 113.5 | 31.8 | 790 1 PLPG | plasmin (EC 3.4.21 |
| 25 | 113.5 | 31.8 | 4548 1 JS0657 | apoptobin(a) (EC |
| 26 | 113 | 31.7 | 810 2 B30848 | plasmin (EC 3.4.21 |
| 27 | 112 | 31.4 | 728 2 A35644 | hepatocyte growth |
| 28 | 111.5 | 31.2 | 603 2 S28941 | coagulation factor |
| 29 | 111 | 31.1 | 812 1 PLBO | plasmin (EC 3.4.21 |

| | | | | |
|----|-------|------|--------------|--------------------|
| 30 | 109 | 30.5 | 685 1 A48289 | neurotrophic recep |
| 31 | 109 | 30.5 | 711 1 A47136 | macrophage-stimula |
| 32 | 108 | 30.3 | 625 1 TBBO | thrombin (EC 3.4.2 |
| 33 | 107.5 | 30.1 | 123 2 B61545 | plasmin (EC 3.4.21 |
| 34 | 106.5 | 29.8 | 434 1 A35005 | u-plasminogen acti |
| 35 | 105.5 | 29.6 | 460 2 B61545 | plasmin (EC 3.4.21 |
| 36 | 105 | 29.4 | 622 1 TBHU | thrombin (EC 3.4.2 |
| 37 | 105 | 29.4 | 710 1 I51283 | hepatocyte growth |
| 38 | 104.5 | 29.3 | 812 1 PLMS | plasmin (EC 3.4.21 |
| 39 | 104 | 29.1 | 560 1 J4795 | plasma hyaluronan- |
| 40 | 101.5 | 28.4 | 120 2 B61545 | plasmin (EC 3.4.21 |
| 41 | 100.5 | 28.2 | 455 2 A61545 | plasmin (EC 3.4.21 |
| 42 | 98.5 | 27.6 | 593 2 S45281 | coagulation factor |
| 43 | 97 | 27.2 | 617 2 SI0511 | thrombin (EC 3.4.2 |
| 44 | 95.5 | 26.8 | 89 2 A60140 | plasmin (EC 3.4.21 |
| 45 | 95 | 26.6 | 618 2 A35827 | thrombin (EC 3.4.2 |

ALIGNMENTS

RESULT 1

UKHUT

t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human

N/Alternate names: t-PA; tissue plasminogen activator

C/Species: Homo sapiens (man)

C/Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text change 08-Dec-2000

C/Accession: A94004; A23529; J70562; A93293; S02125; A91343; A93951; A91322; A54645; I60

R/NY, T.; Eligh, F.; Lund, B.

Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984

A/Title: The structure of the human tissue-type plasminogen activator gene: correlation

A/Reference number: A94004; MUID:84298137; PMID:6089198

A/Accession: A94004

A/Molecule type: DNA

A/Residues: 1-562 <NVT>

A/Cross-references: GB:I00141

A/Note: The codon given for residue 93 (ACC) is inconsistent with the authors' translati.

R/Friener Degen, S.J.; Rajput, B.; Retch, E.

J. Biol. Chem. 261, 6972-6985, 1986

A/Title: The human tissue plasminogen activator gene.

A/Reference number: A23529; MUID:86196143; PMID:3009482

A/Accession: A23529

A/Molecule type: DNA

A/Residues: 1-562 <DEG>

A/Cross-references: GB:K03021; NID:G339817; PIDN:AA98809.1; PID:G339818

R/Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuoka, S.; Higashio, K.

Agric. Biol. Chem. 55, 1225-1232, 1991

A/Title: Purification and characterization of tissue plasminogen activator secreted by h

A/Reference number: J70562; MUID:91291340; PMID:1368681

A/Accession: J70562

A/Molecule type: mRNA

A/Residues: 31-562 <ITR>

A/Cross-references: DDBJ:D01096; NID:G220128; PIDN:BA00881.1; PID:G441174

A/Experimental source: embryonic lung fibroblast IMR-90 cells

A/Note: Part of this sequence, including the amino end of the mature protein, was confir

R/Pennica, D.; Holmes, W.E.; Kohr, W.T.; Hartins, R.N.; Vohar, G.A.; Ward, C.A.; Bennett

Nature 301, 214-221, 1983

A/Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche

A/Reference number: A93293; MUID:83115262; PMID:6337343

A/Accession: A93293

A/Molecule type: mRNA

A/Residues: 1-562 <PEN>

A/Cross-references: GB:I00141

A/Experimental source: melanoma cells

R/Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.

Nucleic Acids Res. 16, 5695, 1988

A/Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe

A/Reference number: S02125; MUID:88262579; PMID:3133640

A/Accession: S02125

A/Status: translation not shown

A/Molecule type: mRNA

A/Residues: 1-562 <SAS>

A/Cross-references: EMBL:X07393; NID:G37243; PIDN:CMA0302.1; PID:G37244

A:Experimental source: fetal lung cells
R:Kakitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Ma
EBBS Lett. 189, 145-149, 1985
A:Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen
A:Reference number: A91343; MUID:85285620; PMID:3896853
A:Accession: A91343
A:Molecule type: mRNA
A:Residues: 1-38, 'G', 86-433, 'E', 435-562 <KAG>
A:Experimental source: Detroit 562 cells; ATCC 138
R:Erdlund, T.; Ny, T.; Randy, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
A:Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen acti
A:Reference number: A93951; MUID:83169566; PMID:6572897
A:Accession: A93951
A:Molecule type: mRNA
A:Residues: 251-358 <EDL>
A:Experimental source: melanoma cells
R:Pohl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H.
Biochemistry 23, 3701-3707, 1984
A:Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived am
differences.
A:Reference number: A90488; MUID:85000468; PMID:6433976
A:Contents: annotation; melanoma cells, partial sequence of residues 36-562, active and
R:Pohl, G.; Kaplan, L.; Sinarsson, M.; Wallen, P.; Jornvall, H.
FEBS Lett. 168, 29-32, 1984
A:Title: Differences between uterine and melanoma forms of tissue plasminogen activator.
A:Reference number: A91322; MUID:84158956; PMID:6538514
A:Accession: A91322
A:Molecule type: protein
A:Residues: 33-45;311-320 <POH>
A:Experimental source: uterus
A:Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln
R:van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.
J. Biol. Chem. 261, 14214-14218, 1986
A:Reference number: A97567; MUID:87033611; PMID:3021732
A:Contents: annotation; fibrin binding site
R:Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Enger
EMBO J. 5, 3525-3530, 1986
A:Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen ac
A:Reference number: A97568; MUID:87161761; PMID:3030730
A:Contents: annotation; fibrin binding site
R:Dodd, I.; Nunn, B.; Robinson, J.H.
Thromb. Haemost. 59, 523-528, 1988
A:Title: Isolation, identification and pharmacokinetic properties of human tissue-type F
A:Reference number: A60902; MUID:89044681; PMID:3142086
A:Contents: annotation; novel forms of expressed recombinant t-PA
R:Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emteage, J.S.; Opdenakker, G.;
Mol. Biol. Med. 3, 279-292, 1986
A:Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its exp
A:Reference number: A54645; MUID:86284200; PMID:3090401
A:Accession: A54645
A:Molecule type: mRNA
A:Residues: 1-562 <HAR>
A:Cross-references: GB:M15518; NID:9190031; PIDN:AAA60111.1; PID:9190032
A:Note: parts of this sequence were confirmed by peptide sequencing
R:Reddy, V.B.; Garramone, A.J.; Sasek, H.; Wei, C.
DNA 6, 461-472, 1987
A:Title: Expression of human uterine tissue-type plasminogen activator in mouse cells us
A:Reference number: I60110; MUID:88054470; PMID:2824147
A:Accession: I60110
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-562 <RES>
A:Cross-references: GB:M18182; NID:9340176; PIDN:AAA36800.1; PID:9340177
R:Fisher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.
J. Biol. Chem. 260, 11223-11230, 1985
A:Title: Isolation and characterization of the human tissue-type plasminogen activator s
A:Reference number: I55232; MUID:85289338; PMID:3161893
A:Accession: I55232
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-36 <RE2>
A:Cross-references: GB:M11890; NID:9339837; PIDN:AAA61213.1; PID:9339839

C:Comment: Cleavage by plasmin or trypsin produces two chains held together by a single
C:Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond. It
C:Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat.
C:Genetics:
A:Gene: GDB:PLAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-8p12
A:Introns: 24/3; 39/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3; 510/
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F:1-17/DNA: signal sequence
F:24-32/DNA: signal sequence #status predicted <PRO>
F:33-562/Product: t-plasminogen activator #status predicted <PRO>
F:33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
F:41-78/DNA: fibronectin type I repeat homology <1F1>
F:86-119/DNA: EGF homology <EGF>
F:127-208/DNA: kringle homology <KR1>
F:215-296/DNA: kringle homology <KR2>
F:311-562/Product: t-plasminogen activator chain B #status experimental <BCH>
F:311-556/DNA: trypsin homology <TRY>
F:41-78-69-97-91-108-110-119-127-208-148-190-179-203-215-226-236-278-267-291-299-4
F:152-483/Binding site: carbohydrate (asn) (covalent) #status experimental
F:219/Binding site: carbohydrate (asn) (covalent) (partial) #status experimental
F:310-311/Cleavage site: Arg-116 (plasmin, trypsin) #status experimental
F:357-406/Active site: His, Asp #status predicted
F:513/Active site: Ser #status experimental

Query Match 40.9%; Score 146; DB 1; Length 562;
Best local Similarity 43.5%; Pred. No. 7.4e-09;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CPMDGNHLVREDQTSAPGRCLNWLDAQSGLASAPVS-----GAGNCGNCPDE 51
DB 127 CYPDGISTRICTWSTASGAECTNW--NSSALAKPFYSGRRDARLGLGNHYCNRPDR 184
QY 52 DPRGPMCVV 60
DB 185 DSK-PWCVV 192

RESULT 2
A35029
C:Plasminogen activator (BC 3.4.21.68) precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
A:Accession: A35029; A31597
R:Feng, P.; Ohlsson, M.; Ny, T.
J. Biol. Chem. 265, 2022-2027, 1990
A:Title: The structure of the TARA-less rat tissue-type plasminogen activator gene. Spec
A:Reference number: A35029; MUID:90130448; PMID:105315
A:Accession: A35029
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-559 <FEN>
A:Cross-references: GB:M1197; NID:9207429; PIDN:AAA42261.1; PID:9207431; GB:J05226
R:Ny, T.; Leonardsson, G.; Heueh, A.J.W.
DNA 7, 671-677, 1988
A:Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activato
A:Reference number: A31597; MUID:89170114; PMID:3148445
A:Accession: A31597
A:Molecule type: mRNA
A:Residues: 1-379, 'K', 381-559 <NYT>
A:Cross-references: GB:M23697; NID:9530159; PIDN:AAA41812.1; PID:9530160
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-17/DNA: signal sequence #status predicted <PRO>
F:18-29/DNA: propeptide #status predicted <PRO>
F:30-559/Product: t-plasminogen activator #status predicted <MAT>
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F:38-75/DNA: fibronectin type I repeat homology <1F1>
F:83-116/DNA: EGF homology <EGF>
F:124-205/DNA: kringle homology <KR1>
F:213-294/DNA: kringle homology <KR2>

F:309-559/Domain: t-Plasminogen activator chain B #status predicted <BCH>
F:309-553/Domain: t-Plasminogen activator chain B #status predicted <BCH>
F:36-68/66-75/83-94/88-105/107-116/124-205/145-187/176-200/213-294/234-276/265-289/297-4
F:149/481/Binding site: carboxylate (Asn) (covalent) #status predicted
F:308-309/Cleavage site: Arg-Ile (Plasmin, trypsin) #status predicted
F:355/404/510/Active site: His, Asp, Ser #status predicted

Query Match 39.5%; Score 141; DB 1; Length 559;
Best Local Similarity 42.0%; Pred. No. 2.7e-08;
Matches 29; Conservative 6; Mismatches 22; Indels 12; Gaps 3;

Qy 1 CFMDNCHLYREDQTPAPGRCINM-----LDAQSGLASPVSGAGNHSYCRNPDE 51
Db 124 CFEGGITYRGWSTAEAGACTNM--NSSALQKPYKAPRRPAIKLGHNHYCRNPDR 181

Qy 52 DPGPWCYV 60
Db 182 DPK-PWCYV 189

RESULT 3

A:Accession: A46688
A:Description: growth factor activator (EC 3.4.21.-) precursor [validated] - human
C:Species: Homo sapiens (man)
C:Date: 21-Sep-1993 #sequence_revision 25-Aug-1995 #text_change 08-Dec-2000
C:Accession: A46688
R:Myazawa, K.; Shimamura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.
J. Biol. Chem. 268, 10024-10028, 1993
A:Title: Molecular cloning and sequence analysis of the cDNA for a human serine protease
d coagulation factor XII
A:Reference number: A46688; MUID:93252878; PMID:7683665
A:Accession: A46688
A:Molecule type: mRNA
A:Residues: 1-655 <M1>
A:Cross-references: DDBJ:D14012; NID:g219680; PIDD:BA03113.1; PID:g219681
A:Experimental source: Liver (mRNA); serum (protein)
A:Note: Sequence extracted from NCBI backbone (NCBI:11227, NCBIP:131228)
A:Note: parts of the sequence, including the amino ends of the heavy and light chains, c
C:Genetics: GDB:HGFA; HGFA; HGFA
A:Cross-references: GDB:9954514
A:Map position: 4p16-4p16

C:Function: A:Description: activates hepatocyte growth factor by specific proteolytic cleavage
A:Pathway: tissue repair and regeneration
C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C:Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase
F:1-34/Domain: signal sequence #status predicted <SIG>
F:108-148/Domain: fibronectin type II repeat homology <1F2>
F:164-197/Domain: EGF homology <EG1>
F:202-237/Domain: fibronectin type I repeat homology <1F1>
F:245-278/Domain: EGF homology <EG2>
F:286-367/Domain: kringle homology <KR3>
F:373-407/Domain: hepatocyte growth factor activator light chain #status experimental
F:408-655/Domain: hepatocyte growth factor activator heavy chain #status experimental
F:408-641/Domain: trypsin homology <TRY>
F:408-48/290/468/492/546/Binding site: carboxylate (Asn) (covalent) #status predicted
F:164-175/169-186/188-197/202-230/228-237/245-256/250-267/269-278/286-367/307-349/338-36
F:447/497/599/Active site: His, Asp, Ser #status predicted

Query Match 39.5%; Score 141; DB 1; Length 655;
Best Local Similarity 45.6%; Pred. No. 3.2e-08;
Matches 31; Conservative 6; Mismatches 21; Indels 10; Gaps 3;

Qy 1 CFMDNCHLYREDQTPAPGRCINM-----LDAQSGLASPVSGAGNHSYCRNPDE 52
Db 286 CFEGGITYRGWSTAEAGACTNM--NSSALQKPYKAPRRPAIKLGHNHYCRNPDR 344

Qy 53 DPGPWCYV 60
Db 345 ER-PWCYV 351

RESULT 4

I38098
t-Plasminogen activator precursor, inactive endothelial splice form - human
N:Alternate names: tissue plasminogen activator
C:Species: Homo sapiens (man)
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
C:Accession: I38098; S01678
R:Siebert, P.D.; Fong, K.
Nucleic Acids Res. 18, 1086, 1990
A:Title: Variant tissue type plasminogen activator (PLAT) cDNA obtained from human endoc

A:Reference number: I38098; MUID:90192128; PMID:1963145
A:Accession: I38098
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-291 <S1E>
A:Cross-references: EMBL:X13097; NID:g35282; PIDD:CAA31489.1; PID:g35283
C:Comment: For the main splice form, see FIR:OKHOT. This form probably does not have pro

C:Genetics: A:Gene: GDB:PLAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-8p12
A:Insertions: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1, 268/2
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: alternative splicing; fibrinolysis; glycoprotein; kringle
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-37/Domain: propeptide #status predicted <PRO>
F:33-251/Product: t-Plasminogen activator, inactive endothelial splice form #status pred
F:41-78/Domain: fibronectin type I repeat homology <1FA>
F:86-119/Domain: EGF homology <EGF>
F:127-208/Domain: kringle homology <KR1>
F:215-291/Domain: kringle homology #status atypical <KR2>
F:41-78/69-78, 86-97, 91-108, 110-119, 127-208, 148-190, 179-203/Diulfide bonds: #status pred

Query Match 38.4%; Score 137; DB 2; Length 291;
Best Local Similarity 42.0%; Pred. No. 4e-08;
Matches 29; Conservative 5; Mismatches 23; Indels 12; Gaps 3;

Qy 1 CFMDNCHLYREDQTPAPGRCINM-----LDAQSGLASPVSGAGNHSYCRNPDE 51
Db 127 CFEGGITYRGWSTAEAGACTNM--NSSALQKPYKAPRRPAIKLGHNHYCRNPDR 184

Qy 52 DPGPWCYV 60
Db 185 DSK-PWCYV 192

Qy 1 CFMDNCHLYREDQTPAPGRCINM-----LDAQSGLASPVSGAGNHSYCRNPDE 51
Db 127 CFEGGITYRGWSTAEAGACTNM--NSSALQKPYKAPRRPAIKLGHNHYCRNPDR 184

Qy 52 DPGPWCYV 60
Db 185 DSK-PWCYV 192

Qy 52 DPGPWCYV 60
Db 185 DSK-PWCYV 192

Qy 52 DPGPWCYV 60
Db 185 DSK-PWCYV 192

Qy 52 DPGPWCYV 60
Db 185 DSK-PWCYV 192

Qy 52 DPGPWCYV 60
Db 185 DSK-PWCYV 192

Qy 52 DPGPWCYV 60
Db 185 DSK-PWCYV 192

Qy 52 DPGPWCYV 60
Db 185 DSK-PWCYV 192

Qy 52 DPGPWCYV 60
Db 185 DSK-PWCYV 192

Qy 52 DPGPWCYV 60
Db 185 DSK-PWCYV 192

```

C-Superfamily fibinase plasminogen activator; EGF homology; fibronectin type I repeat h
C-Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-29/Domain: propeptide #status predicted <PRO>
F:30-559/Product: t-plasminogen activator #status predicted <MAT>
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F:38-75/Domain: fibronectin type I repeat homology <1FI>
F:83-116/Domain: EGF homology <EGF>
F:124-205/Domain: kringle homology <KR1>
F:213-294/Domain: kringle homology <KR2>
F:309-553/Product: t-plasminogen activator chain B #status predicted <BGH>
F:309-553/Domain: t-plasminogen activator #status predicted <TRY>
F:38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-
F:149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
F:355,404,510/Active site: His, Asp, Ser #status predicted

Query Match          38.1%; Score 136; DB 1; Length 559;
Best Local Similarity 40.6%; Pred. No. 1e-07;
Matches 28; Conservative 8; Mismatches 21; Indels 12; Gaps 3;

Oy      1 CFWDNGHLVREDQTSAPGLRCLNWLDAOSGL-----GAGNHSYGRNPDE 51
Db      124 CFEEOCTITRGWSTWSGAECINW--NSSVLTKYPNARNPNAIKLGIDGNHYCRNPDR 181
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Oy      52 DPGKPCVV 60
Db      182 DLK-PWCIV 189
        |::|||::|||

RESULT 6
t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat
N.Alternate names: tissue plasminogen activator
C.Species: Desmodus rotundus (common vampire bat)
C.Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C.Accession: J05059
R.Kiretzechmar, J.; Heendler, B.; Langer, G.; Boldol, W.; Briningmann, P.; Alagon, A.; Dor
Gene 105; 229-237, 1991
A.Title: The plasminogen activator family from the salivary gland of the vampire bat Desm
A.Reference number: J050597, MUID:92039036, PMID:1937019
A.Accession: J05059
A.Molecule type: mRNA
A.Residues: 1-431 <KRA>
A.Cross-references: GB:M63989; NID:G166076; PIDN:AAA31594.1; PID:G166077
C.Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C.Keywords: fibrinolysis; glycoprotein; hydrolyase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:32-36/Domain: propeptide #status predicted <PRO>
F:37-431/Product: plasminogen activator Beta #status predicted <PLA>
F:41-74/Domain: EGF homology <EGF>
F:82-163/Domain: kringle homology <KR3>
F:180-425/Domain: tryptsin homology <TRY>
F:41-52,46-63,65-74,82-163,103-145,134-158,168-299,211-227,219-288,313-388/Disulfide bon
F:119,332/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:179-180/Cleavage site: His-Ser (plasmin) #status predicted
F:326,225,382/Active site: His, Asp, Ser #status predicted
F:345-361,378-406/Disulfide bonds: #status predicted

Query Match          37.5%; Score 134; DB 2; Length 431;
Best Local Similarity 38.6%; Pred. No. 1.3e-07;
Matches 27; Conservative 9; Mismatches 20; Indels 14; Gaps 3;

Oy      1 CFWDNGHLVREDQTSAPGLRCLNWLDAOSGL-----ASAPYGACGNHSYCRNPD 50
Db      82 CYKOGGVTVRGWSTWSGAECINW--NSNILTRRTYNGRSRAITLIGIDGNHYCRNPD 138
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Oy      51 EDPRGPCVV 60
Db      139 NNRSK-PWCIV 147
        |::|||::|||

RESULT 7

```

```

t-plasminogen activator (BC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
A:Accession: A34369
C:Species: Megaderma lyra
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
I:Accession: A34369
R:Garrett, S.J., Duong, L.T., Diehl, R.E., York, J.D.; Hare, T.R.; Register, R.B.; Jacot
J. Biol. Chem. 264, 17947-17952, 1989
A>Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasmiin
A:Reference number: A34369; MUID:90036867; PMID:2509450
A:Accession: A34369
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-477 <GAR>
A:Cross-references: GB:J05082; NID:g166080; PIDN:AAA31593.1; PID:g166081
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F.1-21/Domain: signal sequence #status predicted <SIG>
F.22-36/Domain: propeptide #status predicted <PRO>
F.37-47/Product: plasminogen activator #status predicted <PLA>
F.42-79/Domain: fibronectin type I repeat homology <1FA>
F.87-120/Domain: EGF homology <EGF>
F.128-209/Domain: kringle homology <KRK>
F.226-471/Domain: trypsin homology <TRY>
F.422-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F.185,398/Binding site: carbohydrate (Asn) (covalent) #status predicted
F.225-226/Cleavage site: His-Ser (plasma) #status predicted
F.272,321,428/Active site: His, Asp, Ser #status predicted

Query Match          37.5%; Score 134; DB 1; Length 477;
Best Local Similarity 38.6%; Pred. No. 1.5e-07;
Matches 27; Conservative 9; Mismatches 20; Indels 14; Gaps 3;

QY      1 CFWDNGHLVREDQTSPPAGRLCLNWLDAGSL-----ASAPVSGAGNSHYCRNPD 50
        | . : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db       128 CYKDDGVTVRGWSTSSGSAQCINM--NSNILTRRTYNGRRSDAITGLGNHNYCRNPD 184
        | . : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      51 EDPRGPWCYV 60
        : : : : :
Db       185 NNSK-FMCYV 193

RESULT 8
J:0598
t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0598
R:Kiretzschmar, U.S.; Haendler, B.; Langer, G.; Boisdol, W.; Bringmann, P.; Alagon, A.; Don
Gene 105, 229-237, 1991
A>Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A:Reference number: JS0597; MUID:92039036; PMID:1997019
A:Accession: JS0598
A:Molecule type: mRNA
A:Residues: 1-477 <KRA>
A:Cross-references: GB:M63988; NID:g166074; PIDN:AAA31593.1; PID:g166075
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F.1-21/Domain: signal sequence #status predicted <SIG>
F.22-36/Domain: propeptide #status predicted <PRO>
F.37-47/Product: plasminogen activator alpha-2 #status predicted <PLA>
F.42-79/Domain: fibronectin type I repeat homology <1FA>
F.87-120/Domain: EGF homology <EGF>
F.128-209/Domain: kringle homology <KRK>
F.226-471/Domain: trypsin homology <TRY>
F.422-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F.185,398/Binding site: carbohydrate (Asn) (covalent) #status predicted
F.225-226/Cleavage site: His-Ser (plasma) #status predicted
F.272,321,428/Active site: His, Asp, Ser #status predicted

```


A:Reference number: A51255; PDB:1KDU
 A:Contents: annotation, conformation and disulfide bond assignments by (1)H-NMR, residue
 R1L1, X.; Smith, R.A.G.; Dobson, C.M.
 Biochemistry 31, 9562-9571, 1992
 A>Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain B
 A:Reference number: A44375; MUID:93003110; PMID:1327118
 A:Contents: annotation, conformation and disulfide bond assignments by (1)H-NMR
 R:Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettesheim, D.G.; Mazur, A.P.; Olejniczak,
 submitted to the Brookhaven Protein Data Bank, January 1994
 A:Reference number: A66822; PDB:1URK
 A:Contents: annotation, conformation and disulfide bond assignments by (1)H-NMR, residue
 R:Srreaggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.;
 submitted to the Brookhaven Protein Data Bank, July 1995
 A:Reference number: A66058; PDB:1LMM
 A:Contents: annotation, X-ray crystallography, 2.5 angstroms, residues 168-175,179-426
 C:Comment: This enzyme is found in urine in a high molecular mass form, consisting of A
 C:Comment: urokinase-type plasminogen activator proteolytically activates plasminogen, a
 C:GeneticS:
 A:Gene: GDB:PLAU
 A:Cross-references: GDB:119497; OMIM:191840
 A:Map position: 10q24-10q24
 A:Insertions: 19/3; 23/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/2; 373/3
 C:Function:
 A:Description: proteolytically activates plasminogen
 A:Pathway: fibrinolysis
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C:Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine protease
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-431/Product: urokinase-type plasminogen activator, single chain form #status predic
 F:21-177/Product: urokinase-type plasminogen activator chain A #status experimental <MR>
 F:31-62/Domain: EGF homology <RGF>
 F:70-151/Domain: kringle homology <KRG>
 F:168-177/Product: urokinase-type plasminogen activator chain A1 #status experimental <M
 F:179-431/Product: urokinase-type plasminogen activator chain B #status experimental <M
 F:179-419/Domain: trypsin homology <TRY>
 F:31-33,33-51,53-62,70-151,91-133,132-146,168-299,209-225,217-288,313-382,345-361,372-40
 F:38/Binding site: carbohydrate (Thr) (covalent) #status predicted
 F:178-179/Cleavage site: Lys-Ile (plasmin) #status experimental
 F:224,275,376/Active site: His, Asp, Ser #status experimental
 F:332/Binding site: carbohydrate (Asn) (covalent) #status experimental
 Query Match 36.7%; Score 131; DB 1; Length 431;
 Best Local Similarity 40.3%; Pred. No. 2,9e-07;
 Matches 27; Conservative 7; Mismatches 25; Indels 8; Gaps 2;
 QY 1 CFPDNGHLYREDQTSFAPGLACLMWDA-----QSGIASAPVSGAGNSYCNPDDEP 53
 Db 70 CFEONGHFRKGASTDTWGRPLPMSATVLTQTHAHRSDALQGLGKHNYCNPD-NR 128
 QY 54 RGPWCYV 60
 Db 129 RRPWCYV 135
 RESULT 11
 UKPG
 u-plasminogen activator (EC 3.4.21.73) precursor - pig
 N:Alternate names: uPA
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
 C:Accession: A00932
 R:Nagamine, Y.; Pearson, D.; Altun, M.S.; Reich, E.
 Nucleic Acids Res. 12, 9525-9541, 1984
 A>Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
 A:Reference number: A00932; MUID:85087954; PMID:6096832
 A:Accession: A00932
 A:Molecule type: DNA
 A:Residues: 1-240, 'H', 242-442 <NAG1>
 A:Experimental source: kidney cell line LLC-PK1
 R:Nagamine, Y.
 submitted to the Protein Sequence Database, December 1986
 A:Reference number: A37566
 A:Contents: annotation, correction to residue 241

C:GeneticS:
 A:Insertions: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F:31-64/Domain: EGF homology <EGF>
 F:72-153/Domain: kringle homology <KRG>
 F:190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F:190-430/Domain: trypsin homology <TRY>
 F:152/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:179-310,220-236,228-299,324-393,356-372,383-411/Disulfide bonds: #status predicted
 F:235,286,387/Active site: His, Asp, Ser #status predicted
 Query Match 36.4%; Score 130; DB 1; Length 442;
 Best Local Similarity 44.8%; Pred. No. 3,9e-07;
 Matches 30; Conservative 4; Mismatches 25; Indels 8; Gaps 3;
 QY 1 CFPDNGHLYREDQTSFAPGLACLMWDA-----ASAPVSGAGNSYCNPDDEP 53
 Db 72 CFEONGHFRKGASTDTWGRPLPMSATVLTQTHAHRSDALQGLGKHNYCNPD-NQ 130
 QY 54 RGPWCYV 60
 Db 131 RRPWCYV 137
 RESULT 12
 JCS061
 macrophage-stimulating protein 1 precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 16-Jun-2000
 C:Accession: JCS061
 R:Ohnishi, K.; Iwama, A.; Matsuno, K.; Ezaki, T.; Sakamoto, O.; Hamaguchi, I.; Takasu, N
 Biochem. Biophys. Res. Commun. 227, 273-280, 1996
 A>Title: Molecular cloning of rat macrophage-stimulating protein and its involvement in
 A:Reference number: JCS061; MUID:97011126; PMID:8858136
 A:Accession: JCS061
 A:Molecule type: mRNA
 A:Residues: 1-716 <OHS>
 A:Cross-references: EMBL:X95096; NID:G1669718; PID:CA64473.1; PID:G1669719
 C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
 C:Keywords: duplication; glycoprotein; growth factor; kringle
 F:1-11/Domain: signal sequence #status predicted <SIG>
 F:32-488,489-716/Product: macrophage-stimulating protein 1 #status predicted <MAT>
 F:32-488/Domain: macrophage-stimulating protein 1 alpha chain #status predicted <ACH>
 F:110-186/Domain: kringle homology <KR11>
 F:191-268/Domain: kringle homology <KR12>
 F:292-370/Domain: kringle homology <KR13>
 F:379-457/Domain: kringle homology <KR14>
 F:489-716/Domain: macrophage-stimulating protein 1 beta chain #status predicted <BCH>
 F:489-709/Domain: trypsin homology <TRY>
 F:72,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted
 Query Match 35.9%; Score 128; DB 1; Length 716;
 Best Local Similarity 41.9%; Pred. No. 1,1e-06;
 Matches 26; Conservative 3; Mismatches 29; Indels 4; Gaps 2;
 QY 1 CFPDNGHLYREDQTSFAPGLACLMW-----LDAQSLASAPVSGAGNSYCNPDDEP 57
 Db 110 CIMPNGHSTRGVARTADGLPCQAMSRFPNDHKYTPTRKNGL-ENPCRNPDGPRGPW 168
 QY 58 CY 59
 Db 169 CY 170
 RESULT 13
 JN0560
 u-plasminogen activator (EC 3.4.21.73) precursor - bovine
 N:Alternate names: uPA
 C:Species: Bos primigenius taurus (cattle)

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
 C/Accession: JN0560
 R/Kreisetschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schlenker, W.D.
 Gene 125, 177-183, 1993
 A>Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and induc
 A/Reference number: JN0560; PMID:93216119; PMID:8385052
 A/Accession: JN0560
 A/Molecule type: mRNA
 A/Residues: 1-433 <KRA>
 A/Cross-references: GB:U03546; NID:G163800; PIDN:AAA51419.1; PID:G163801
 C/Superfamily: urokinase-type plasminogen activator; EGF homology; kringie homology; try
 C/Keywords: glycoprotein; heterodimer; hydrolase; kringie; serine proteinase
 F:1-80/Domain: signal sequence #status predicted <SIG>
 F:21-179/Product: plasminogen activator chain A #status predicted <MA1>
 F:21-179/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F:33-64/Domain: EGF homology <EGF>
 F:72-153/Domain: kringie homology <KR>
 F:181-433/Product: plasminogen activator chain B #status predicted <MA2>
 F:181-421/Domain: trypsin homology <TRY>
 F:170-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
 F:126,277,378/Active site: His, Asp, Ser #status predicted

Query Match 35.6%; Score 127; DB 1; Length 433;
 Best Local Similarity 41.2%; Pred. No. 8,3e-07;
 Matches 28; Conservative 6; Mismatches 24; Indels 10; Gaps 3;

OY 1 CFMDNGHLVREDQTSPPAGLRCLNWDASGL-----ASAPVSGAGNHSYCRNPDED 52
 DB 72 CYQNGHSHYKGNKRNDRDLSGRPCILAW-DSPTVLKMYAHNRSDALQGLGSKINCRNDP-N 129
 OY 53 PRGPMCVY 60
 DB 130 QRPMCVY 137

RESULT 14
 US0600
 t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
 N/Alternate names: tissue plasminogen activator
 C/Species: Desmodus rotundus (common vampire bat)
 C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C/Accession: JN0600
 R/Kreisetschmar, J.; Haendler, B.; Langer, G.; Boldol, W.; Bringmann, P.; Alagon, A.; Dor
 Gene 105, 229-237, 1991
 A>Title: The plasminogen activator family from the salivary gland of the vampire bat Des
 A/Reference number: JN0597; PMID:92039036; PMID:1937019
 A/Accession: JN0600
 A/Molecule type: mRNA
 A/Residues: 1-394 <KRA>
 A/Cross-references: GB:M63990; NID:G166078; PIDN:AAA31595.1; PID:G166079
 A>Note: the authors translated the codon ATC for residue 75 as Thr
 C/Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringie; serine proteinase
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:12-36/Domain: propeptide #status predicted <PRO>
 F:37-394/Product: plasminogen activator gamma #status predicted <PLA>
 F:45-126/Domain: kringie homology <KR>
 F:143-388/Domain: trypsin homology <TRY>
 F:45-126,66-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide bonds:
 F:142-143/Cleavage site: His-Ser (plasmin) #status predicted
 F:189,238,345/Active site: His, Asp, Ser #status predicted
 F:315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 35.3%; Score 126; DB 2; Length 394;
 Best Local Similarity 37.3%; Pred. No. 9,8e-07;
 Matches 25; Conservative 8; Mismatches 26; Indels 8; Gaps 2;

OY 1 CFMDNGHLVREDQTSPPAGLRCLNWDASGLASAPVSGAGNHSYCRNPDED 53
 DB 45 CYQNGHSHYKGNKRNDRDLSGRPCILAW-DSPTVLKMYAHNRSDALQGLGSKINCRNDP 104
 OY 54 PRGPMCVY 60
 DB 130 QRPMCVY 137

DB 105 K-PMCVY 110

RESULT 15
 A40332
 macrophage-stimulating protein 1 precursor - mouse
 N/Alternate names: hepatocyte growth factor-1-like protein
 C/Species: Mus musculus (house mouse)
 C/Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 18-Jun-1999
 A/Accession: A40332; B40332
 R/Degen, S.J.F.; Stuart, L.A.; Han, S.; Jamison, C.S.
 Biochemistry 30, 9781-9791, 1991
 A>Title: Characterization of the mouse cDNA and gene coding for a hepatocyte growth fact
 A/Reference number: A40332; PMID:20020217; PMID:1832957
 A/Accession: A40332
 A/Molecule type: DNA
 A/Residues: 1-716 <DEG>
 A/Cross-references: GB:M74180; NID:G193831; PIDN:AAA50166.1; PID:G193832
 A/Accession: B40332
 A/Molecule type: mRNA
 A/Residues: 1-18, 'P', 20-716 <DEG2>
 A/Cross-references: GB:M74181; NID:G193833; PIDN:AAA50167.1; PID:G193834
 C/Genetics:
 A:Insertions: 18/1, 67/2, 105/1, 143/2, 189/1, 229/2, 269/1, 334/2, 378/1, 412/2, 458/1, 47
 C/Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C/Superfamily: hepatocyte growth factor; kringie homology; trypsin homology
 C/Keywords: duplication; glycoprotein; growth factor; kringie
 F:1-31/Domain: signal sequence #status predicted <SIG>
 F:19-488,489-716/Product: macrophage-stimulating protein 1 #status experimental <MAT>
 F:19-488/Domain: alpha chain #status experimental <ACH>
 F:110-186/Domain: kringie homology <KR1>
 F:191-268/Domain: kringie homology <KR2>
 F:292-370/Domain: kringie homology <KR3>
 F:379-457/Domain: kringie homology <KR4>
 F:484-711/Domain: beta chain #status experimental <BCH>
 F:489-709/Domain: trypsin homology <TRY>
 F:72,173,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 35.3%; Score 126; DB 1; Length 716;
 Best Local Similarity 41.9%; Pred. No. 1,8e-06;
 Matches 26; Conservative 3; Mismatches 29; Indels 4; Gaps 2;

OY 1 CFMDNGHLVREDQTSPPAGLRCLNWDASGLASAPVSGAGNHSYCRNPDEPRGPW 57
 DB 110 CYMDNGVSYRGVARTAGGLPCQAMSRFPMDKRYTPFANGL-ENFGNPDGDPGPM 168
 OY 58 CY 59
 DB 169 CY 170

Search completed: April 7, 2003, 09:19:20
 Job time : 5.86486 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:24 ; Search time 2.74428 Seconds
(without alignments)
906.823 Million cell updates/sec

Title: US-10-057-951-2_COPY_25_84

Perfect score: 357

Sequence: 1 CFWDNGHLYREDQTSPPAGL.....GNHSYCRNPDEDPGRGWCYV 60

Scoring table: BIOSUM62

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 154 | 43.1 | 566 | 1 TPA_BOVIN | Q28198 bos taurus |
| 2 | 146 | 40.9 | 562 | 1 TPA_HUMAN | P00750 homo sapien |
| 3 | 146 | 40.9 | 653 | 1 HGFA_MOUSE | O90708 mus musculu |
| 4 | 141 | 39.5 | 559 | 1 TPA_RAT | P19637 rattus norv |
| 5 | 141 | 39.5 | 655 | 1 HGFA_HUMAN | O04756 homo sapien |
| 6 | 136 | 38.1 | 559 | 1 TPA_MOUSE | P11214 mus musculu |
| 7 | 134 | 37.5 | 431 | 1 URTB_DESRO | P98121 desmodus ro |
| 8 | 134 | 37.5 | 477 | 1 URT2_DESRO | P15638 desmodus ro |
| 9 | 132 | 37.0 | 433 | 1 UROK_PAPCY | P16227 papio cynoc |
| 10 | 131 | 36.7 | 431 | 1 UROK_HUMAN | P00749 homo sapien |
| 11 | 130 | 36.4 | 442 | 1 UROK_PIG | P04185 sus scrofa |
| 12 | 127 | 35.6 | 433 | 1 UROK_BOVIN | O05589 bos taurus |
| 13 | 126 | 35.3 | 394 | 1 URTG_DESRO | P49150 desmodus ro |
| 14 | 126 | 35.3 | 716 | 1 HGFL_MOUSE | P26928 mus musculu |
| 15 | 123 | 34.5 | 169 | 1 PLMN_RAT | O01177 rattus norv |
| 16 | 121 | 33.9 | 810 | 1 PLMN_HUMAN | P00747 homo sapien |
| 17 | 119 | 33.3 | 433 | 1 UROK_MOUSE | P06869 mus musculu |
| 18 | 119 | 33.3 | 728 | 1 HGF_HUMAN | P14210 homo sapien |
| 19 | 118 | 33.1 | 432 | 1 UROK_RAT | P29598 rattus norv |
| 20 | 118 | 33.1 | 728 | 1 HGF_MOUSE | O08048 mus musculu |
| 21 | 118 | 33.1 | 810 | 1 PLMN_ERIEU | Q29485 erinaceus e |
| 22 | 115 | 32.2 | 477 | 1 URT1_DESRO | P89119 desmodus ro |
| 23 | 113.5 | 31.8 | 333 | 1 PLMN_CANFA | P80009 canis fami |
| 24 | 113.5 | 31.8 | 790 | 1 PLMN_PIG | P06867 sus scrofa |
| 25 | 113.5 | 31.8 | 4548 | 1 APOA_HUMAN | P08519 homo sapien |
| 26 | 113 | 31.7 | 810 | 1 PLMN_MACMU | P12545 macaca mula |
| 27 | 112 | 31.4 | 728 | 1 HGF_RAT | P17945 rattus norv |
| 28 | 111.5 | 31.2 | 603 | 1 FA12_CAVPO | O04962 cavia porce |
| 29 | 111 | 31.1 | 812 | 1 PLMN_BOVIN | P06868 bos taurus |
| 30 | 110 | 30.8 | 473 | 1 KREM_MOUSE | O29443 mus musculu |
| 31 | 110 | 30.8 | 473 | 1 KREM_RAT | O29444 rattus norv |
| 32 | 110 | 30.8 | 475 | 1 KREM_HUMAN | O96mu8 homo sapien |
| 33 | 109 | 30.5 | 711 | 1 HGFL_HUMAN | P26927 homo sapien |

| | | | | | |
|----|-------|------|------|--------------|--------------------|
| 34 | 108 | 30.3 | 625 | 1 THRB_BOVIN | P00735 bos taurus |
| 35 | 106.5 | 29.8 | 434 | 1 UROK_CHICK | P15120 gallus gall |
| 36 | 105 | 29.4 | 622 | 1 THRB_HUMAN | P00734 homo sapien |
| 37 | 104.5 | 29.3 | 812 | 1 PLMN_MOUSE | P20921 mus musculu |
| 38 | 98.5 | 27.6 | 593 | 1 FA12_BOVIN | P98140 bos taurus |
| 39 | 97.5 | 27.3 | 343 | 1 PLMN_SHEEP | P81286 ovis aries |
| 40 | 97 | 27.2 | 617 | 1 THRB_RAT | P18292 rattus norv |
| 41 | 95.5 | 26.8 | 338 | 1 PLMN_HORSE | P18010 equus cabal |
| 42 | 95 | 26.6 | 618 | 1 THRB_MOUSE | P19221 mus musculu |
| 43 | 94.5 | 26.5 | 325 | 1 PLMN_PETMA | P33574 petromyzon |
| 44 | 94.5 | 26.5 | 615 | 1 FA12_HUMAN | P00748 homo sapien |
| 45 | 93.5 | 26.2 | 1420 | 1 APOA_MACMU | P14117 macaca mula |

ALIGNMENTS

```

RESULT 1
TPA_BOVIN          STANDARD;          PRT;          566 AA.
ID                ID
AC  Q28198;
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE  (t-PA) (t-Plasminogen activator).
GN  PLAT.
OS  Bos taurus (Bovine).
OC  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC  Bovidae; Bovinae; Bos.
OX  NCBI_TaxID=9913;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Kidney;
RA  Ravn P., Berglund L., Petersen T.E.;
RT  "Cloning and characterization of the bovine plasminogen activators uPA
RT  and tPA."
RL  Int. Dairy J. 5:605-617(1995).
CC  -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
CC  TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC  CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC  ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC  MANY OTHER PHYSIOPATHOLOGICAL EVENTS.
CC  -!- CATALYTIC ACTIVITY: Specific cleavage of Arg|-Val bond in
CC  plasminogen to form plasmin.
CC  -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC  BOND.
CC  -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC  -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC  PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC  ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC  -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC  CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC  -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC  -!- SIMILARITY: CONTRAINS 1 EGF-LIKE DOMAIN.
CC  -!- SIMILARITY: CONTRAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC  -!- SIMILARITY: CONTRAINS 2 KRINGLE DOMAINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X85800; CAA59795.1; -.
DR  HSSP; P00750; 1RTP.
DR  MEROPS; S01.232; -.
DR  InterPro; IPR001314; Chymotrypsin.
DR  InterPro; IPR000561; EGF-like.
DR  InterPro; IPR000083; Fibrinctn1.

```

```

DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; Kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR Prodom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS0186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Plasma; Kringie; EGF-like domain; Repeat; Signal.
KW Plasma; Kringie; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 21
FT PROPEP 22 33
FT CHAIN 34 566
FT CHAIN 34 314
FT CHAIN 315 566
FT CHAIN 315 566
FT DOMAIN 40 82
FT DOMAIN 83 121
FT DOMAIN 128 209
FT DOMAIN 219 300
FT DOMAIN 315 566
FT ACT_SITE 361 410
FT ACT_SITE 410 410
FT ACT_SITE 517 517
FT DISULFID 72 72
FT DISULFID 70 79
FT DISULFID 87 98
FT DISULFID 92 109
FT DISULFID 111 120
FT DISULFID 128 209
FT DISULFID 149 191
FT DISULFID 180 204
FT DISULFID 219 300
FT DISULFID 240 282
FT DISULFID 271 295
FT DISULFID 303 434
FT DISULFID 346 362
FT DISULFID 354 423
FT DISULFID 448 523
FT DISULFID 480 496
FT CARBOHYD 513 541
FT CARBOHYD 153 153
FT CARBOHYD 487 487
FT SEQUENCE 566 AA; 63701 MW; 2EB6BB4E32276C3 CRC64;

Query Match 43.1%; Score 154; DB 1; Length 566;
Best Local Similarity 44.9%; Pred. No. 2e-10;
Matches 31; Conservative 6; Mismatches 20; Indels 12; Gaps 3;

```

```

TPA_HUMAN
ID TPA_HUMAN STANDARD; PRT; 562 AA.
AC P00750; Q15103.
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.66) (tPA)
DE (t-PA) (t-plasminogen activator) (Alteplase) (reteplase).
GN PLAT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=83115262; PubMed=6337343;
RA Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vohar G.A.,
RA Ward C.A., Bennett W.F., Velterton E., Seeburg P.H., Heyneker H.L.,
RA Goeddel D.V., Collen D.;
RT "Cloning and expression of human tissue-type plasminogen activator
RT cDNA in E. coli.";
RL Nature 301:214-221(1983).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=88262579; PubMed=3133640;
RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niva M.;
RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA
RT from human fetal lung cells.";
RL Nucleic Acids Res. 16:5695-5695(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88054470; PubMed=2824147;
RA Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Galli J.,
RA Heising N.;
RT "Expression of human uterine tissue-type plasminogen activator in
RT mouse cells using BPV vectors.";
RL DNA 6:461-472(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=86196143; PubMed=3009482;
RA Fieznex Degen S.J., Rajput B., Reich E.;
RT "The human tissue plasminogen activator gene.";
RL J Biol. Chem. 261:6972-6985(1986).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=84298137; PubMed=6089198;
RA Ny T., Elgh F., Lund B.;
RT "The structure of the human tissue-type plasminogen activator gene:
RT correlation of intron and exon structures to functional and
RT structural domains.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=86284200; PubMed=3090401;
RA Harris T.J., Patel T., Marston F.A., Little S., Emage J.S.,
RA Opdenakker G., Volckaert G., Rombauts W., Billiau A., Somer P.;
RT "Cloning of cDNA coding for human tissue-type plasminogen activator
RT and its expression in Escherichia coli.";
RL Mol. Biol. Med. 3:279-292(1986).
RN [7]
RP SEQUENCE OF 212-361 FROM N.A.
RX MEDLINE=83169656; PubMed=6572897;
RA Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,
RA Josephson S.;
RT "Isolation of cDNA sequences coding for a part of human tissue
RT plasminogen activator.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).
RN [8]
RP SEQUENCE OF 1-36 FROM N.A.
RX MEDLINE=85289338; PubMed=3161893;
RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,

```

RA Schleuning W.-D.;
 RT "Isolation and characterization of the human tissue-type plasminogen
 RT activator structural gene including its 5' flanking region.";
 RL J. Biol. Chem. 260:11223-11230 (1985).
 RN [9]
 RP SEQUENCE OF 31-562 FROM N.A.
 RX MEDLINE=91291340; PubMed=1366881;
 RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;
 RT "Purification and characterization of tissue plasminogen activator
 RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells.";
 RL Agric. Biol. Chem. 55:1225-1232 (1991).
 RN [10]
 RP SEQUENCE OF 36-562.
 RC TISSUE-Melanoma;
 RX MEDLINE=8500468; PubMed=6433976;
 RA Pohl G., Kaelinstroem M., Bergsdorf N., Wallen P., Joernvall H.;
 RT "Tissue plasminogen activator: peptide analyses confirm an indirectly
 RT derived amino acid sequence, identify the active site serine residue,
 RT establish glycosylation sites, and localize variant differences.";
 RL Biochemistry 23:3701-3707 (1984).
 RN [11]
 RP SEQUENCE OF 33-52 AND 311-330.
 RC TISSUE-Melanoma;
 RX MEDLINE=83209620; PubMed=6682760;
 RA Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;
 RT "Purification and characterization of a melanoma cell plasminogen
 RT activator.";
 RL Eur. J. Biochem. 132:681-686 (1983).
 RN [12]
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).
 RC TISSUE=umbilical vein;
 RX MEDLINE=90192129; PubMed=2107528;
 RA Siebert P.D., Fong K.;
 RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
 RT human endothelial cells.";
 RL Nucleic Acids Res. 18:1086-1086 (1990).
 RN [13]
 RP STRUCTURE OF CARBOHYDRATES.
 RX MEDLINE=90092112; PubMed=2513186;
 RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;
 RT "Carbohydrate structure of recombinant human uterine tissue
 RT plasminogen activator expressed in mouse epithelial cells.";
 RL Eur. J. Biochem. 186:273-286 (1989).
 RN [14]
 RP CARBOHYDRATE-LINKAGE SITE THR-96.
 RX MEDLINE=91159408; PubMed=1900431;
 RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;
 RT "Tissue plasminogen activator has an O-linked fucose attached to
 RT threonine-61 in the epidermal growth factor domain.";
 RL Biochemistry 30:2311-2314 (1991).
 RN [15]
 RP DISULFIDE BONDS IN KRINGLE 2.
 RX MEDLINE=91244765; PubMed=1645336;
 RA Vlahos C.J., Wilhelm O.G., Hassell T., Taskunas S.R., Bang N.U.;
 RT "Disulfide pairing of the recombinant kringle-2 domain of tissue
 RT plasminogen activator produced in *Bescherichia coli*.";
 RL J. Biol. Chem. 266:10070-10072 (1991).
 RN [16]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=96200985; PubMed=8613982;
 RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
 RA Bode W.;
 RT "The 2.3 Å crystal structure of the catalytic domain of recombinant
 RT two-chain human tissue-type plasminogen activator.";
 RL J. Mol. Biol. 258:117-135 (1996).
 RN [17]
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=97449126; PubMed=9305622;
 RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,
 RA Bode W.;
 RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray
 RT crystal structure of single-chain human tPA.";
 RL EMBO J. 16:4797-4805 (1997).

RN [18]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.
 RX MEDLINE=92118803; PubMed=1310033;
 RA de Vos A., Ulsch M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,
 RA Westbrook M.L., Kosiakof A.A.;
 RT "Crystal structure of the kringle 2 domain of tissue plasminogen
 RT activator at 2.4-Å resolution.";
 RL Biochemistry 31:270-279 (1992).
 RN [19]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=90122799; PubMed=2558718;
 RA Byeon I.-J.L., Kelley R.F., Lillias M.;
 RT "1H NMR structural characterization of a recombinant kringle 2 domain
 RT from human tissue-type plasminogen activator.";
 RL Biochemistry 28:9350-9360 (1989).
 RN [20]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=91200042; PubMed=1901789;
 RA Byeon I.-J.L., Kelley R.F., Lillias M.;
 RT "Kringle-2 domain of the tissue-type plasminogen activator. 1H-NMR
 RT assignments and secondary structure.";
 RL Eur. J. Biochem. 197:155-165 (1991).
 RN [21]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=92106329; PubMed=1762144;
 RA Byeon I.-J.L., Lillias M.;
 RT "Solution structure of the tissue-type plasminogen activator kringle
 RT 2 domain complexed to 6-aminohexanoic acid an antifibrinolytic
 RT drug.";
 RL J. Mol. Biol. 222:1035-1051 (1991).
 RN [22]
 RP STRUCTURE BY NMR OF 38-85.
 RX MEDLINE=92292163; PubMed=1602484;
 RA Downing A.K., Driscoll P.C., Harvey T.S., Dudgeon T.J., Smith B.O.,
 RA Baron M., Campbell I.D.;
 RT "Solution structure of the fibrin binding finger domain of
 RT tissue-type plasminogen activator determined by 1H nuclear magnetic
 RT resonance.";
 RL J. Mol. Biol. 225:821-833 (1992).
 RN [23]
 RP STRUCTURE BY NMR OF 36-126.
 RX MEDLINE=96027104; PubMed=7582899;
 RA Smith B.O., Downing A.K., Driscoll P.C., Dudgeon T.J., Campbell I.D.;
 RT "The solution structure and backbone dynamics of the fibronectin type
 RT I and epidermal growth factor-like pair of modules of tissue-type
 RT plasminogen activator.";
 RL Structure 3:823-833 (1995).
 CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
 CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
 CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
 CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
 CC MANY OTHER PHYSIOLOGICAL EVENTS.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg1-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: BINDS TO FIBRIN WITH HIGH AFFINITY. THIS INTERACTION

Query Match 40.9%; Score 146; DB 1; Length 562;
 Best Local Similarity 43.5%; Pred. No. 1.6e-09;
 Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

| | | | |
|----|-----|--|-----|
| QY | 1 | CFWNGHLYREDQSPAPRGCLNWLDAQSLAPVS-----GAGNHSYCRPDE | 51 |
| DB | 127 | CYEDQGISYRTGWTSTASGACTNW--NSSALQKPYSGRRPDAIRLGLGNNHNCRRPDR | 164 |
| QY | 52 | DPRGWCYV 60 | |
| | | : | |
| DB | 185 | DSK-PWCYV 192 | |

RESULT 3
 HGFA_MOUSE ID_HGFA_MOUSE STANDARD; PRT; 653 AA.
 Q9R098; Q9JUK4;

16-OCT-2001 (Rel. 40, Created)
 15-OCT-2001 (Rel. 40, Last sequence update)
 15-JUN-2002 (Rel. 41, Last annotation update)
 Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF activator) (HGFA).
 HGFAC.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OK NCBI_Taxid=10090;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RA Itoh H., Kataoka H., Koono H.;
 RL "Mouse hepatocyte growth factor activator.";
 RT Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 [2]
 RN SEQUENCE FROM N.A.
 RA van Adelsberg J.S., Sehgal S., Kukes A., Brady C., Barasch J., Yang J., Huan Y.;
 RT "Activation of HGF by endogenous HGF activator is required for mesangial kidney morphogenesis in vitro.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1 FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM (BY SIMILARITY).
 CC -1 SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A DISULFIDE BOND (BY SIMILARITY).
 CC -1 SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM (BY SIMILARITY).
 CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1 SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -1 SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
 CC -1 SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.
 CC -1 SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 CC EMBL; AF099017; AAF02489.1; -;
 DR EMBL; AF224724; AAF34712.1; -;
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.228; -;
 DR MGD; MG11859281; Hgfac.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000742; EGF 2.
 DR InterPro: IPR000562; FN Type II.
 DR InterPro: IPR000083; F1Fnctn1.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF_2.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; tryptain; 1.
 DR PRINTS; PRO0722; CHYMOTRYPsin.
 DR PRINTS; PRO0013; FNTYPEII.
 DR PRINTS; PRO0018; KRINGLE.
 DR PRODOM; PD000395; Kringle; 1.
 DR PRODOM; PD000995; FN_Type_II; 1.
 DR SMART; SM00181; EGF_2.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRYp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 1.

DR PROSITE; PS01253; FIBRONECTIN 1; 1.
 DR PROSITE; PS00023; FIBRONECTIN 2; 1.
 DR PROSITE; PS00021; KRINGLE 1; 1.
 DR PROSITE; PS00070; KRINGLE 2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydroxylase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
 KM EGF-like domain; Repeat; Zymogen.
 FT SIGNAL 1 29
 FT PROPEP 30 369
 FT CHAIN 370 405
 FT CHAIN 406 653
 FT DOMAIN 105 145
 FT DOMAIN 157 195
 FT DOMAIN 197 237
 FT DOMAIN 238 276
 FT DOMAIN 283 364
 FT ACT_SITE 406 653
 FT ACT_SITE 445 445
 FT ACT_SITE 495 495
 FT ACT_SITE 596 596
 FT DISULFID 105 130
 FT DISULFID 119 145
 FT DISULFID 161 172
 FT DISULFID 166 183
 FT DISULFID 185 194
 FT DISULFID 199 227
 FT DISULFID 225 234
 FT DISULFID 242 253
 FT DISULFID 247 264
 FT DISULFID 266 275
 FT DISULFID 283 364
 FT DISULFID 304 346
 FT DISULFID 335 359
 FT DISULFID 392 519
 FT DISULFID 430 446
 FT DISULFID 438 508
 FT DISULFID 533 602
 FT DISULFID 565 581
 FT DISULFID 592 620
 FT CARBOHYD 39 39
 FT CARBOHYD 47 47
 FT CARBOHYD 63 63
 FT CARBOHYD 287 287
 FT CARBOHYD 466 466
 FT CARBOHYD 544 544
 FT CONFLICT 164 164
 SQ SEQUENCE 653 AA; 70567 MW; 88B4B20255DF7FDC CRC64;

 Query Match 40.9%; Score 146; DB 1; Length 653;
 Best Local Similarity 47.1%; Pred. No. 1.9e-09;
 Matches 32; Conservative 7; Mismatches 19; Indels 10; Gaps 3;

 QY 1 CFMNGCHLYREDQTSPPAPGLCLNW-----LDAOSGLASAPVSGAGNHSYCRNDED 52
 Db 283 CFLNGTEYRVAVASTASGLAMNSDLVQELHVDV-VAAAVLGLGPHAYCRNPKD 341
 QY 53 PRGPGCVY 60
 Db 342 ER-FMCTV 348

 RESULT 4
 ID TPA_RAT STANDARD; PRT; 559 AA.
 AC P19637;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)

ID HGFA HUMAN STANDARD; PRT; 655 AA.
AC 004756; 014726;
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-1994 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF activator) (HGFA).
GN HGFA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC TISSUE=Liver, and Serum;
RX MEDLINE=93252878; PubMed=763665;
RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y., Kitamura N.;
RT "Molecular cloning and sequence analysis of the cDNA for a human serine protease responsible for activation of hepatocyte growth factor. Structural similarity of the protease precursor to blood coagulation factor XII.";
RT J. Biol. Chem. 268:10024-10028 (1993).
RN [2]
RP SEQUENCE OF 40-655 FROM N.A.
RA Zhao S., Odell C.;
RL Submitted (Feb-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM.
CC -1- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A DISULFIDE BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM.
CC -1- TISSUE SPECIFICITY: LIVER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D14012; BAA03113.1; -;
DR EMBL; Z69923; CA93803.1; -;
DR PIR; A46688; A46688.
DR HSRP; P00763; IDPO.
DR MEROPS; S01.228; -;
DR Genew; HSNCA494; HGFA.
DR MIM; 604552; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF-2.
DR InterPro; IPR000562; FN Type II.
DR InterPro; IPR000083; p130catl.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00039; fn2; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FMTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000995; FN_Type_II; 1.

DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR HydroLase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 30
FT PROPEP 31 372
FT CHAIN 373 407
FT CHAIN 408 655
FT CHAIN 655 655
FT DOMAIN 108 148
FT DOMAIN 160 198
FT DOMAIN 200 240
FT DOMAIN 241 279
FT DOMAIN 286 367
FT DOMAIN 408 655
FT ACT_SITE 447 497
FT ACT_SITE 497 497
FT ACT_SITE 598 598
FT ACT_SITE 108 133
FT DISULFID 122 148
FT DISULFID 164 175
FT DISULFID 169 186
FT DISULFID 188 197
FT DISULFID 202 230
FT DISULFID 228 237
FT DISULFID 245 256
FT DISULFID 250 267
FT DISULFID 269 278
FT DISULFID 286 367
FT DISULFID 307 349
FT DISULFID 338 362
FT DISULFID 394 521
FT DISULFID 432 448
FT DISULFID 440 510
FT DISULFID 535 604
FT DISULFID 567 583
FT DISULFID 594 622
FT CARBOHYD 468 468
FT CARBOHYD 290 290
FT CARBOHYD 468 468
FT CARBOHYD 492 492
FT CARBOHYD 546 546
FT CONFLICT 644 644
SQ SEQUENCE 655 AA; 70681 MW; 2CF72P1E1B862ED7 CRC64;
Query Match 39.5%; Score 141; DB 1; Length 655;
Best Local Similarity 45.6%; Pred. No. 7.3e-05;
Matches 31; Conservative 6; Mismatches 21; Indels 10; Gaps 3;
QY 1 CFMDNGHLYVEDQTSAPQIRCLNM-----LDAOSGLASAPVSGAGNHSYCRNPDED 52
DB 286 CFLNGTGTVGAVSTASGLSCLANNSDLVQELHVS-VGAALTLGLGPHAYCRNPEDND 344
QY 53 PRGPWCYV 60
DB 345 ER-PWCYV 351
RESULT 6
TPA_MOUSE

| ID | TPA_MOUSE | STANDARD; | PRT; | 559 AA. |
|----|--|-----------------------------------|------|---------|
| AC | P11214 | | | |
| DT | 01-JUL-1989 | (Rel. 11, Created) | | |
| DT | 01-JUL-1989 | (Rel. 11, Last sequence update) | | |
| DT | 15-JUN-2002 | (Rel. 41, Last annotation update) | | |
| DE | Tissue-type plasminogen activator precursor (BC 3.4.21.68) (tPA) | | | |
| DE | (t-PA) (t-plasminogen activator). | | | |
| GN | PLAT. | | | |
| OS | Mus musculus (Mouse). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |
| OX | NCBI_Taxid=10090; | | | |
| RP | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=88087303; PubMed=2826484; | | | |
| RA | Rickles R.J., Darrow A.L., Strickland S.; | | | |
| RT | "Molecular cloning of complementary DNA to mouse tissue plasminogen | | | |
| RT | activator mRNA and its expression during F9 teratocarcinoma cell | | | |
| RT | differentiation." | | | |
| RL | J. Biol. Chem. 263:1563-1569 (1988). | | | |
| CC | -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN | | | |
| CC | TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY | | | |
| CC | CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT | | | |
| CC | ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND | | | |
| CC | MANY OTHER PHYSIOPATHOLOGICAL EVENTS. | | | |
| CC | -1- CATALYTIC ACTIVITY: Specific cleavage of Arg- -Val bond in | | | |
| CC | plasminogen to form plasmin. | | | |
| CC | -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE | | | |
| CC | BOND. | | | |
| CC | -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR. | | | |
| CC | -1- PFM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER | | | |
| CC | PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER | | | |
| CC | ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA. | | | |
| CC | -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A | | | |
| CC | CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY. | | | |
| CC | -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. | | | |
| CC | -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN. | | | |
| CC | -1- SIMILARITY: CONTAINS 1 FIBRINOCTIN TYPE I DOMAIN. | | | |
| CC | -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS. | | | |
| CC | ----- | | | |
| CC | THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION | | | |
| CC | BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION - | | | |
| CC | THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS | | | |
| CC | USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY | | | |
| CC | MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL | | | |
| CC | ENTITIES REQUIRES A LICENSE AGREEMENT (See http://www.isb-sab.ch/announce/ | | | |
| CC | OR SEND AN EMAIL TO license@isb-sab.ch). | | | |
| CC | ----- | | | |
| DR | EMBL; J03520; AAAA0470.1; - | | | |
| DR | PIR; A29941; A29941. | | | |
| DR | HSSE; P00750; 1A5H. | | | |
| DR | MEROPS; S01.232; - | | | |
| DR | MGD; MGI:97610; Plat. | | | |
| DR | InterPro: IPR001314; Chymotrypsin. | | | |
| DR | InterPro: IPR000561; EGF-like. | | | |
| DR | InterPro: IPR000083; Fibnctn1. | | | |
| DR | InterPro: IPR000001; Kringle. | | | |
| DR | InterPro: IPR001254; Ser protease_Try. | | | |
| DR | Pfam; PF00008; EGF_1. | | | |
| DR | Pfam; PF00039; fn1; 1. | | | |
| DR | Pfam; PF00051; kringle; 2. | | | |
| DR | Pfam; PF00089; trypsin; 1. | | | |
| DR | PRINTS; PR00722; CHYMOTRYPSIN. | | | |
| DR | PRINTS; PR00018; KRINGLE. | | | |
| DR | ProDom; PDO00395; Kringle; 2. | | | |
| DR | SMART; SM00181; EGF_1. | | | |
| DR | SMART; SM00058; FN1; 1. | | | |
| DR | SMART; SM00130; KR; 2. | | | |
| DR | SMART; SM00020; Tryp_Spc; 1. | | | |
| DR | PROSITE; PS00022; EGF_1; 1. | | | |
| DR | PROSITE; PS01186; EGF_2; 1. | | | |
| DR | PROSITE; PS01253; FIBRINOCTIN_1; 1. | | | |
| DR | PROSITE; PS00021; KRINGLE_1; 2. | | | |

| | | |
|---|---|---|
| DR | PROSITE; PS50070; | KRINGLE_2; 2. |
| DR | PROSITE; PSS0240; | TRYPSIN_DOM; 1. |
| DR | PROSITE; P800134; | TRYPSIN_HIS; 1. |
| DR | PROSITE; PS00133; | TRYPSIN_SER; 1. |
| KW | Plasminogen activation; | Hydrolase; Serine protease; Glycoprotein; |
| KW | Plasma; Kringle; | Egfr-like domain; Repeat; Signal. |
| FT | SIGNAL | 1 |
| FT | PROPEP | 18 29 |
| FT | CHAIN | 30 559 |
| FT | CHAIN | 30 308 |
| FT | CHAIN | |
| FT | CHAIN | 309 559 |
| FT | DOMAIN | 36 78 |
| FT | DOMAIN | 79 117 |
| FT | DOMAIN | 124 205 |
| FT | DOMAIN | 213 294 |
| FT | DOMAIN | 309 559 |
| FT | ACT_SITE | 355 355 |
| FT | ACT_SITE | 404 404 |
| FT | ACT_SITE | 510 510 |
| FT | DISULFID | 38 68 |
| FT | DISULFID | 66 75 |
| FT | DISULFID | 83 94 |
| FT | DISULFID | 88 105 |
| FT | DISULFID | 107 116 |
| FT | DISULFID | 124 205 |
| FT | DISULFID | 145 187 |
| FT | DISULFID | 176 200 |
| FT | DISULFID | 213 294 |
| FT | DISULFID | 234 276 |
| FT | DISULFID | 265 289 |
| FT | DISULFID | 297 428 |
| FT | DISULFID | 340 356 |
| FT | DISULFID | 348 417 |
| FT | DISULFID | 442 516 |
| FT | DISULFID | 474 490 |
| FT | DISULFID | 506 534 |
| FT | CARBOHYD | 149 149 |
| FT | CARBOHYD | 481 481 |
| SO | SEQUENCE | 559 AA; 63110 MW; 4ACCE57DC6A282A5 CRC64; |
| <hr/> | | |
| Query Match | | |
| Best Local Similarity 38.1%; Score 136; DB 1; Length 559; | | |
| Matches 28; Conservative 8; Mismatches 21; Indels 12; Gaps 3; | | |
| Qy | 1 CPWDNGHLVREDOGTSPAPGLRCINMLDAOSGIASAPVS-----GAGNHSYCRNPDE 51 | |
| Db | 124 CFEBCGITRYRGWWSMAESAECINN--NSSVSLKPYNAARRPAIKLGLGNHNYCENPDR 181 | |
| Qy | 52 DPRGPWCYV 60 | |
| Db | 182 DLK-PWCYV 189 | |
| <hr/> | | |
| RESULT 7 | | |
| URTB_DESRO | STANDARD; | PRT; 431 AA. |
| ID | URTB_DESRO | |
| AC | P98121; | |
| DT | 01-FEB-1996 (Rel. 33, Created) | |
| DT | 01-FEB-1996 (Rel. 33, Last sequence update) | |
| DT | 15-JUN-2002 (Rel. 41, Last annotation update) | |
| DE | Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSBA beta). | |
| OS | Desmodus rotundus (Vampire bat). | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| OC | Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae; | |
| OC | Desmodontinae; Desmodus. | |
| OX | NCHI_Taxid=9430; | |
| RN | [1] | |
| RP | SEQUENCE FROM N.A. | |
| RS | TISSUE=Salivary gland; | |
| RX | MEDLINE=92039036; PubMed=1937019; | |

RA Kretzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
 RA Alagon A., Donner P., Schleuning W.D.;
 RT "The plasminogen activator family from the salivary gland of the
 RT vampire bat *Desmodus rotundus*: cloning and expression.";
 RL Gene 105:229-237(1991).
 RN (2)
 RP CHARACTERIZATION.
 RX MEDLINE=93393059; PubMed=1309059;
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
 RA Kretzschmar J., Haendler B., Langer G., Baldus B., Wilt W.,
 RA Donner P.;
 RT "Plasminogen activators from the saliva of *Desmodus rotundus* (common
 RT vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
 CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
 CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
 CC AGENT.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).

 DR EMBL: M63989; AAA31594.1; -
 DR HSSP: P98119; 1A51.
 DR MEROPS: S01.239; -
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00008; EGF_1.
 DR Pfam: PF00051; kringle_1.
 DR Pfam: PF00089; trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle_1.
 DR SMART: SM00181; EGF_1.
 DR SMART: SM00130; KR_1.
 DR SMART: SM00020; Tryp_Spc_1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS00070; KRINGLE_2; 1.
 DR PROSITE: PS00240; TRYPsin_DOM_1.
 DR PROSITE: PS00134; TRYPsin_HIS_1.
 DR PROSITE: PS00135; TRYPsin_SER_1.
 KW Plasminogen activation; Hydrolyse; Serine protease; Glycoprotein;
 FT Kringle; EGF-like domain; Signal; Multigene family.
 FT SIGNAL 1 36
 FT CHAIN 37 431
 FT DOMAIN 82 75
 FT DOMAIN 179 163
 FT ACT_SITE 226 431
 FT ACT_SITE 275 275
 FT ACT_SITE 382 382
 FT DISULFID 41 52
 FT DISULFID 46 63
 FT DISULFID 65 74
 FT DISULFID 82 163
 FT DISULFID 103 145
 FT DISULFID 134 158
 FT DISULFID 168 299
 FT DISULFID 211 227
 BY SIMILARITY.

FT DISULFID 219 288 BY SIMILARITY.
 FT DISULFID 313 388 BY SIMILARITY.
 FT DISULFID 345 361 BY SIMILARITY.
 FT DISULFID 378 406 BY SIMILARITY.
 FT CARBOHYD 139 139 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 431 AA; 48221 MW; 69958675B162CBF CRC64;
 Query Match 37.5%; Score 134; DB 1; Length 431;
 Best Local Similarity 38.6%; Pred. No. 3.1e-08;
 Matches 27; Conservative 9; Mismatches 20; Indels 14; Gaps 3;
 QY 1 CFPDNGHLTHREDTSPAPGRLCLNWDAGSGL-----ASAPVSGAGHSTCYRNP 50
 DB 82 CYDDGVTVRGWSTSTSGAGCINW---NSNLTTRTYNGRSDAITGLGNHNYCRNP 138
 QY 51 EDRGPMCVY 60
 DB 139 NNSK-PWCIV 147
 RESULT 8
 ID URT2 DESRO STANDARD: PRT; 477 AA.
 AC P15638;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSBA
 OS alpha-2) (BAT-PA) (T-plasminogen activator).
 OS *Desmodus rotundus* (Vampire bat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 OC Desmodontinae; *Desmodus*.
 OC NCBI_Taxid=9430;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=salivary gland;
 RX MEDLINE=92039036; PubMed=1937019;
 RA Kretzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
 RA Alagon A., Donner P., Schleuning W.D.;
 RT "The plasminogen activator family from the salivary gland of the
 RT vampire bat *Desmodus rotundus*: cloning and expression.";
 RL Gene 105:229-237(1991).
 RN [2]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=salivary gland;
 RX MEDLINE=90036867; PubMed=2509450;
 RA Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,
 RA Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;
 RT "Isolation, characterization, and cDNA cloning of a vampire bat
 RT salivary plasminogen activator.";
 RL J. Biol. Chem. 264:17947-17952(1989).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=93393059; PubMed=1309059;
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
 RA Kretzschmar J., Haendler B., Langer G., Baldus B., Wilt W.,
 RA Donner P.;
 RT "Plasminogen activators from the saliva of *Desmodus rotundus* (common
 RT vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
 CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
 CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
 CC AGENT.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN
 CC THE PRESENCE OF FIBRIN I.
 CC -1- SUBUNIT: MONOMER.
 CC -1- DOMAIN: THE FIBRINECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,
 CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
 CC STIMULATION OF ACTIVITY.

| | | | |
|----|----------|--|-----|
| CC | -1-* | SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. | |
| CC | -1- | SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN. | |
| CC | -1- | SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN. | |
| CC | -1- | SIMILARITY: CONTAINS 1 KRINGLE DOMAIN. | |
| CC | | ----- | |
| CC | | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). | |
| CC | | ----- | |
| DR | EMBL | M63988; AAA31593.1; - | |
| DR | EMBL | J05082; AAA31596.1; - | |
| DR | PIR | A34369; A34369. | |
| DR | HSSP | P98119; 1A5I. | |
| DR | MEROPS | S01_232; - | |
| DR | InterPro | IPR001314; Chymotrypsin. | |
| DR | InterPro | IPR000561; EGF-like. | |
| DR | InterPro | IPR000083; Fibronectn. | |
| DR | InterPro | IPR000001; Kringle. | |
| DR | InterPro | IPR001254; Ser_protease_Try. | |
| DR | Pfam | PF00008; EGF_1. | |
| DR | Pfam | PF00039; fn1; 1. | |
| DR | Pfam | PF00051; kringle; 1. | |
| DR | Pfam | PF00089; trypsin; 1. | |
| DR | PRINTS | PR00722; CHIMOTRYPSIN. | |
| DR | PRINTS | PR00018; KRINGLE. | |
| DR | ProDom | PD000395; Kringle; 1. | |
| DR | SMART | SM00181; EGF; 1. | |
| DR | SMART | SM00058; FN1; 1. | |
| DR | SMART | SM00130; KR; 1. | |
| DR | SMART | SM00020; TRYD_SPC; 1. | |
| DR | PROSITE | PS00023; EGF_1; 1. | |
| DR | PROSITE | PS01186; EGF_2; 1. | |
| DR | PROSITE | PS01253; FIBRONECTIN_1; 1. | |
| DR | PROSITE | PS00021; KRINGLE_1; 1. | |
| DR | PROSITE | PS00270; KRINGLE_2; 1. | |
| DR | PROSITE | PS50240; TRYPSIN_DOM; 1. | |
| DR | PROSITE | PS00134; TRYPSIN_HIS; 1. | |
| DR | PROSITE | PS00135; TRYPSIN_SER; 1. | |
| KM | | plasmogen activation; Hydrolase; Serine protease; Glycoprotein; kringle; EGF-like domain; Signal; Multigene family. | |
| KM | | ----- | |
| FT | SIGNAL | 1 | 36 |
| FT | CHAIN | 37 | 477 |
| FT | DOMAIN | 40 | 82 |
| FT | DOMAIN | 83 | 121 |
| FT | DOMAIN | 128 | 209 |
| FT | DOMAIN | 225 | 477 |
| FT | ACT_SITE | 272 | 321 |
| FT | ACT_SITE | 321 | 321 |
| FT | ACT_SITE | 428 | 428 |
| FT | DISULFID | 42 | 72 |
| FT | DISULFID | 70 | 79 |
| FT | DISULFID | 87 | 98 |
| FT | DISULFID | 92 | 109 |
| FT | DISULFID | 111 | 120 |
| FT | DISULFID | 128 | 209 |
| FT | DISULFID | 149 | 191 |
| FT | DISULFID | 180 | 204 |
| FT | DISULFID | 214 | 345 |
| FT | DISULFID | 257 | 273 |
| FT | DISULFID | 265 | 334 |
| FT | DISULFID | 359 | 434 |
| FT | DISULFID | 391 | 407 |
| FT | DISULFID | 424 | 452 |
| FT | CARBOHYD | 185 | 398 |
| FT | CARBOHYD | 398 | 403 |
| FT | CONFLICT | 403 | 403 |
| FT | CONFLICT | 417 | 417 |
| FT | CONFLICT | 417 | 417 |
| FT | CONFLICT | 435 | 435 |
| FT | SEQUENCE | 477 AA; 53719 MW; 17486555COE5077C CnC64; | |

```

Query Match: 37.5%; Score 134; DB 1; Length 477;
Best Local Similarity 38.6%; Pred. No. 3,4e-08;
Matches 27; Conservative 9; Mismatches 20; Indels 14; Gaps 3;

Dy 1 CFMDNGHLTYREDQTDSPAGRLCLMWLDNQSL-----ASAFVSGAGNHSICRND 50
Db 128 CYKQGGVTRGTWSTSSGACINW---NSNLTLRRTYGRGRSDATLTGLGNHNYCRND 184
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 51 EDPGPGVCYV 60
      : : : : :
Db 185 NNSK-PCYV 193

RESULT 9
UROK_PAPCY STANDARD; PRT; 433 AA.
AC P16227;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
DE PLAU.
OC Papio cynocephalus (Yellow baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Papio.
OC NCBI_TaxID=9556;
RN [1]
RP SOURCE FROM N.A.
RC TISSUE=Thoracic aorta;
RX MEDLINE=90287734; PubMed=2113276;
RA Au Y.P.T., Wang T.W., Clowes A.W.;
RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
RT plasminogen activator.";
RL Nucleic Acids Res. 18:3411-3411(1990).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X51935; CAA36200.1; -.
DR PIR; S14687; UKBAY.
DR HSSP; P00749; 11MW.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRODOM; PD000395; kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS01066; EGF_1; 1.
DR PROSITE; PS01066; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.

```

DR PROSITE; PS50070; KRINGLE 2; 1.
DR PROSITE; PS50240; TRYPSIN DOM; 1.
DR PROSITE; PS00134; TRYPSIN HIS; 1.
DR PROSITE; PS00135; TRYPSIN SER; 1.
KM PLASMINOGEN activation; Hydrolyase; Serine protease; Glycoprotein;
KM Kringle; EGF-like domain; zymogen; signal.
FT SIGNAL 1 20
FT CHAIN 1 433
FT CHAIN 21 176
FT CHAIN 155 176
FT CHAIN 178 433
FT CHAIN 26 62
FT DOMAIN 69 150
FT DOMAIN 151 177
FT DOMAIN 178 433
FT DISULFID 30 38
FT DISULFID 32 50
FT DISULFID 52 61
FT DISULFID 167 298
FT DISULFID 208 224
FT DISULFID 216 287
FT DISULFID 315 384
FT DISULFID 347 363
FT DISULFID 374 402
FT ACT_SITE 223 223
FT ACT_SITE 274 274
FT ACT_SITE 378 378
FT CARBOHYD 324 324
SQ SEQUENCE 433 AA; 48595 MW; 816D22DFEDDC8792 CRC64;
Query Match 37.0%; Score 132; DB 1; Length 433;
Best local Similarity 40.3%; Pred. No. 5.3e-08;
Matches 27; Conservative 7; Mismatches 25; Indels 8; Gaps 2;
QY 1 CFWDNGHLYREDQTSPPAPGRLWLDL-----QSGLSAPVSGAGNHYCRNEDDP 53
DB 69 CYEGNCHFYRGKASDTWGRSCLAMNSATVYQQTYYAHRSDALQLGKHNYCRND-NR 127
QY 54 RGPWCYV 60
DB 128 RRMWCYV 134
RESULT 10
ID UROK_HUMAN STANDARD; PRT; 431 AA.
AC P00749; O15844; O1618; O969W6;
DT 21-JUL-1986 (Rel. 01, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
DE (U-plasminogen activator).
GN PLAU.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RX MEDLINE=85215647; PubMed=2987867;
RA Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blasi F.;
RT "The human urokinase-plasminogen activator gene and its promoter";
RL Nucleic Acids Res. 13:2759-2771(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX Holmes W.E., Pennic D., Blaber M., Rey M.W., Guenzler W.A.,
RA Steffens G.J., Heynaker H.L.;
RT "Cloning and expression of the gene for pro-urokinase in Escherichia
col.";
RL Biotechnology 3:923-929(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=85056954; PubMed=2415429;

RA Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Aritamura H.,
RA Nishida M., Suyama T.;
RT "Molecular cloning of cDNA coding for human preprourokinase";
RL Gene 36:183-188(1985).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=85203359; PubMed=3888571;
RA Jacobs P., Cravador A., Lortiau R., Brockly F., Colau B., Chuchana P.,
RA van Elsen A., Herzog A., Boilen A.;
RT "Molecular cloning, sequencing, and expression in Escherichia coli of
human preprourokinase cDNA";
RL DNA 4:139-146(1985).
RN [5]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC Tissue=Lung;
RA Strausberg R.;
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE OF 66-431 FROM N.A.
RX MEDLINE=84272706; PubMed=589620;
RA Verde P., Stopelli M.P., Galetti P., di Nocera P., Blasi F.;
RT "Identification and primary sequence of an unspliced human urokinase
poly(A)+ RNA";
RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
RN [8]
RP SEQUENCE OF 21-177.
RX MEDLINE=83055084; PubMed=6754569;
RA Gunzler W.A., Steffens G.J., Oetting F., Kim S.-M.A., Frankus E.,
RA Flohe L.;
RT "The primary structure of high molecular mass urokinase from human
urine. The complete amino acid sequence of the A chain";
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
RN [9]
RP SEQUENCE OF 156-176 AND 179-224.
RX MEDLINE=83003608; PubMed=6749491;
RA Schaller J., Nick H., Rickli E.E., Gillesen D., Lergier W.,
RA Studer R.O.;
RT "Human low-molecular-weight urinary urokinase. Partial
characterization and preliminary sequence data of the two polypeptide
chains";
RL Eur. J. Biochem. 125:251-257(1982).
RN [10]
RP SEQUENCE OF 158-410.
RX MEDLINE=83055099; PubMed=6754572;
RA Steffens G.J., Gunzler W.A., Oetting F., Frankus E., Flohe L.;
RT "The complete amino acid sequence of low molecular mass urokinase
from human urine";
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=9600858; PubMed=8591045;
RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
RA Dobson C.M., Stuart D.I., Jones E.Y.;
RT "The crystal structure of the catalytic domain of human
urokinase-type plasminogen activator";
RL Structure 3:681-691(1995).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
RX MEDLINE=20266327; PubMed=10805774;
RA Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,
RA Bode W., Magdalen V., Huber R., Moroder L.;
RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly
selective inhibitors of human urokinase";
RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
RN [13]
RP STRUCTURE BY NMR.
RX MEDLINE=89127526; PubMed=2536903;
RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.;

RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-dimensional NMR.";
 RT Nature 337:579-582(1989).
 RL [14]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=93003110; PubMed=1327118;
 RA Li X., Smith R.A.G., Dobson C.M.;
 RT "Sequential 1H NMR assignments and secondary structure of the kringle domain from urokinase.";
 RL Biochemistry 31:9562-9571(1992).
 RN [15]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=94149701; PubMed=8107091;
 RA Li X., Bokman A.M., Lijonas M., Smith R.A.G., Dobson C.M.;
 RT "Solution structure of the kringle domain from urokinase-type plasminogen activator.";
 RL J. Mol. Biol. 235:1548-1559(1994).
 RN [16]
 RP VARIANT LEU-141.
 RX MEDLINE=96186279; PubMed=8652631;
 RA Yoshimoto M., Uehiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K., Sawasaki Y., Hamada K.;
 RT "Characterization of single chain urokinase-type plasminogen activator with a novel amino-acid substitution in the kringle structure.";
 RL Biochim. Biophys. Acta 1293:83-89(1996).
 RN [17]
 RP VARIANT LEU-141.
 RX MEDLINE=97218551; PubMed=9065988;
 RA Come B., Berczy M., Belin D.;
 RT "Detection of polymorphisms in the human urokinase-type plasminogen activator gene.";
 RL Thromb. Haemost. 77:434-435(1997).
 RN [18]
 RP Thromb. Haemost. 77:434-435(1997).
 RP ERRATUM.
 RA Come B., Berczy M., Belin D.;
 RL Thromb. Haemost. 78:973-973(1997).
 RN [19]
 RP VARIANT LEU-141.
 RX MEDLINE=97337920; PubMed=9194591;
 RA Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Hell W., Creutzburg S., Graeff H., Magdolen V.;
 RT "Molecular analysis of the genes encoding urokinase-type plasminogen activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
 RL Electrophoresis 18:686-689(1997).
 CC -1- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR THERAPY OF THROMBOLYTIC DISORDERS.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.
 CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.
 CC -1- PHARMACEUTICAL: Available under the name Abbocinase (Abbott). Used in Pulmonary Embolism (PE) to initiate fibrinolysis.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation- the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X02419; CA26268.1; -
 DR EMBL; M15476; AAA61253.1; -
 DR EMBL; D00244; BAA00175.1; -
 DR EMBL; D11443; BAA01919.1; -
 DR EMBL; X02760; CAA26535.1; -
 DR EMBL; AF377330; AAK53822.1; -

DR EMBL; BC013575; AAH13575.1; -
 DR EMBL; K03226; AAC97138.1; -
 DR EMBL; K02286; AAA61252.1; -
 DR EMBL; A21571; CAA01559.1; -
 DR EMBL; A18397; CAA01390.1; -
 DR PIR; A00931; UKRH.
 DR PIR; A32974; A32974.
 DR PDB; 1KDU; 31-OCT-93.
 DR PDB; 1LWV; 29-JAN-96.
 DR PDB; 1URK; 08-MAY-95.
 DR PDB; 1EJN; 17-MAY-00.
 DR MEROPS; S01.231; -
 DR GlycoSuiteDB; P00749; -
 DR Genew; HGNC:9052; PLAU.
 DR MIM; 191840; -
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00051; Kringle1; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRODOM; PD000395; Kringle1; 1.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.

Query Match 36.7%; Score 131; DB 1; Length 431;
 Best Local Similarity 40.3%; Pred. No. 6.9e-08;
 Matches 27; Conservative 7; Mismatches 25; Indels 8; Gaps 2;

QY 1 CFWDNCHLYREDQTSPPARGLRCINWLD-----OSGLASAPVSGAGNHCYCRNPDDP 53
 Db 70 CYEGNGHFFYRGKASITDTPGRPCLPMSATVLCQTYHAHNSDALQGLGGRNCRNPD-NR 128
 QY 54 RGPACTV 60
 Db 129 RRPWCYV 135

RESULT 11
 ID _UROK_PIG STANDARD; PRT; 442 AA.
 AC P04185;
 DT 20-MAR-1987 (rel. 04, Created)
 DT 13-AUG-1987 (rel. 05, Last sequence update)
 DT 15-JUN-2002 (rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxId=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=85087954; PubMed=6096832;
 RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
 RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
 RL Nucleic Acids Res. 12:9525-9541(1984).
 RN [2]
 RP REVISION TO 241.
 RA Nagamine Y.;
 RL Submitted (DEC-1986) to the PIR data bank.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X01648; CAA25806.1; -
DR EMBL; X02724; CAA26511.1; -
DR PIR; A00932; UNPG.
DR HSSP; P00749; IKDU.
DR MEROPS; S01.231; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; trypsin; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringleg; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KM Plasminogen activation; Hydrolyase; Signal.
KW Kringleg; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 188 CHAIN A (BY SIMILARITY).
FT CHAIN 130 442 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 189 CONNECTING PEPTIDE.
FT DOMAIN 442 442 SERINE PROTEASE.
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .).
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 119 310 INTERCHAIN (BY SIMILARITY).
FT DISULFID 220 236 BY SIMILARITY.
FT DISULFID 228 299 BY SIMILARITY.
FT DISULFID 324 393 BY SIMILARITY.
FT DISULFID 356 372 BY SIMILARITY.
FT DISULFID 383 411 BY SIMILARITY.
FT ACT_SITE 235 235 CHARGE RELAY SYSTEM.
FT ACT_SITE 286 286 CHARGE RELAY SYSTEM.
FT ACT_SITE 387 387 CHARGE RELAY SYSTEM.
FT ACT_SITE 241 241 Q -> H (IN REF. 1; CAA25806).
FT ACT_SITE 242 242 Q -> H (IN REF. 1; CAA26511).
FT ACT_SITE 288 288 A -> G (IN REF. 1; CAA25806).
FT CONFLICT 288 288
SQ SEQUENCE 442 AA; 49116 MW; E832PCFE501321EB CRC64;

Query Match 36.4%; Score 130; DB 1; Length 442;
Best local Similarity 44.8%; Pred. No. 9.3e-08;
Matches 30; Conservative 4; Mismatches 25; Indels 8; Gaps 3;

QY 1 CFNDNGHLTREDDTSAPPLRLNWLDAQSL-----ASAPVS---GAGNHSYCRANDEDP 53
DB 72 CFEGNHSHYRKANKNTGGRPLTPNSATVTLNTYHARPDALQGLGKRNKYNCRND-NQ 130
QY 54 RGPWCYV 60
DB 131 RRPWCYV 137
RESULT 12
UROK_BOVIN STANDARD; PRT; 433 AA.
ID UROK_BOVIN
AC Q05589; Q8209;

DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UBA)
DE (U-plasminogen activator).
GN PLAU.
OC Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue: Aortic endothelium;
RX MEDLINE=93216119; PubMed=835052;
RA Kretzschmar U., Haendler B., Kojima S., Rifkin D.B.,
RA Schleming W.-D.,
RT "Bovine urokinase-type plasminogen activator and its receptor:
RT cloning and induction by retinoic acid";
RL Gene 125:177-183(1993).
RN [2]
RP SEQUENCE OF 12-433 FROM N.A.
RC Tissue: Kidney;
RA Ravn P., Berglund L., Petersen T.E.,
RT "Cloning and characterization of the bovine plasminogen activators uPA
RT and tPA";
RL Int. Dairy J. 5:605-617(1995).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-Val bond in
CC plasminogen to form plasmin.
CC -1- INDUCTION: BY RETINOIC ACID.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; L03546; AAA51419.1; -
DR EMBL; X83801; CAA39796.1; -
DR PIR; JN0560; JN0560.
DR HSSP; P00749; ILMW.
DR MEROPS; S01.231; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringleg.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR Pfam; PF00051; Kringleg; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringleg; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KM Plasminogen activation; Hydrolyase; Signal.
KW Kringleg; EGF-like domain; Zymogen.
FT SIGNAL 1 20
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 179 CHAIN A (BY SIMILARITY).
FT CHAIN 181 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 180 CONNECTING PEPTIDE.

```

FT DOMAIN 181 433 SERINE PROTEASE.
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CONFLICT 189 189 A -> T (IN REF. 2).
SQ SEQUENCE 433 AA; 48730 MW; 4DE1BD4DA47027A CRC64;

Query Match 35.6%; Score 127; DB 1; Length 433;
Best Local Similarity 41.2%; Pred. No. 2e-07; Mismatches 24; Indels 10; Gaps 3;
Matches 28; Conservative 6;

QY 1 CFWMDGHLVREDQTSPPAGLRCLNMLDAQSGL-----ASAPVSGAGNHSYCRNPDED 52
DB 72 CYQDQGVTVRSTESBSGACQIMWNSNLIIRTYNGRMPEAVKLGUNHNYCRNPDGAS 104
QY 53 RGPWCYV 60
DB 130 RGPWCYV 137

RESULT 13
UTRG DESRO STANDARD; PRT; 394 AA.
AC P49150:
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSPA gamma).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae; Desmodontinae; Desmodus.
OC NCBI_Taxid=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kretschmar J., Haendler B., Langer G., Boidol W., Bringmann P., Alagon A., Donner P., Schlenning W.D.;
RT "The plasminogen activator family from the salivary gland of the vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237 (1991).
RN [2]
RX CHARACTERIZATION; PubMed=1309059;
RX MEDLINE=93393059;
RA Schlenning W.-D., Alagon A., Boidol W., Bringmann P., Petri T., Kretschmar J., Haendler B., Langer G., Baldus B., Witt W., Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403 (1992).
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC AGENT.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.
CC -!- SUBUNIT: MONOMER.
CC -!- SIMILARITY: BELONGS TO PERTINASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M63990; AAA31595.1; -.
DR HSSP; P98119; 1A51.
DR MEROPS; S01.239; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000004; Kringle.
DR InterPro; IPR001254; Ser protease_Try.
DR Pfam; PF00051; kringle_1.
DR Pfam; PF00089; trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle_1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
KW Kringle; Signal; Multigene family.
FT CHAIN 1 36 POTENTIAL.
FT SIGNAL 1 36
FT DOMAIN 37 394 SALIVARY PLASMINOGEN ACTIVATOR GAMMA.
FT ACT_SITE 142 126 KRINGLE.
FT ACT_SITE 189 189 SERINE PROTEASE.
FT ACT_SITE 238 238 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 345 345 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 45 126 BY SIMILARITY.
FT DISULFID 66 108 BY SIMILARITY.
FT DISULFID 97 121 BY SIMILARITY.
FT DISULFID 131 262 BY SIMILARITY.
FT DISULFID 174 190 BY SIMILARITY.
FT DISULFID 182 251 BY SIMILARITY.
FT DISULFID 276 351 BY SIMILARITY.
FT DISULFID 308 324 BY SIMILARITY.
FT DISULFID 341 369 BY SIMILARITY.
FT CARBOHYD 315 315 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 394 AA; 44105 MW; 9CDD6F52F3D81FCD CRC64;

Query Match 35.3%; Score 126; DB 1; Length 394;
Best Local Similarity 37.3%; Pred. No. 2.4e-07; Mismatches 26; Indels 8; Gaps 2;
Matches 25; Conservative 8;

QY 1 CFWMDGHLVREDQTSPPAGLRCLNMLDAQSGLASAPVSGAGNHSYCRNPDED 53
DB 45 CYQDQGVTVRSTESBSGACQIMWNSNLIIRTYNGRMPEAVKLGUNHNYCRNPDGAS 104
QY 54 RGPWCYV 60
DB 105 K-PWCYV 110

RESULT 14
HGFL MOUSE STANDARD; PRT; 716 AA.
ID HGFL_MOUSE
AC P26928;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor-like protein precursor (Macrophage stimulatory protein) (MSP).
DE MST1 OR HGFL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.

```



```
CC -----
DR EMBL; M62832; AAA41884.1; -.
DR PIR; A40522; A40522.
DR HSSP; P00747; 1PMK.
DR MEROPS; S01.233; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF000051; kringle; 2.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00001; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRYPsin_DOM; PARTIAL.
DR PROSITE; PS00134; TRYPsin_HIS; PARTIAL.
DR PROSITE; PS00135; TRYPsin_SER; PARTIAL.
DR Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Repeat.
FT NON_TER 1
FT DOMAIN 1
FT DOMAIN <1 10 KRINGLE 3 (BY SIMILARITY).
FT DOMAIN 34 112 KRINGLE 4 (BY SIMILARITY).
FT DOMAIN 139 >169 KRINGLE 5 (BY SIMILARITY).
FT DISULFID 34 112 BY SIMILARITY.
FT DISULFID 55 95 BY SIMILARITY.
FT DISULFID 83 107 BY SIMILARITY.
FT NON_TER 169
SQ SEQUENCE 169 AA; 18401 MW; 77A54214C9D010C CRC64;

Query Match 34.5%; Score 123; DB 1; Length 169;
Best local Similarity 37.5%; Pred. No. 2, 2e-07;
Matches 24; Conservative 9; Mismatches 25; Indels 6; Gaps 2;

QY 1 CFWDNGHLYREDQSPAPGLRCLNWLDA-----QSGLASAPVSGAGNHSYCNRPDPERG 55
DB 34 CIGGNGKSYKGTSSITWTGKCKQSWMTSPHSHSKTPANFPDSGL-EMNYCNPNPDORG 92
QY 56 PWCY 59
DB 93 PWCF 96
```

Search completed: April 7, 2003, 09:16:57
Job time : 3.74428 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:25 ; Search time 16.0915 Seconds

(without alignments)
768.284 Million cell updates/sec

Title: US-10-057-951-2_COPY_25_84

Perfect score: 357
Sequence: 1 CFWDNGHLTYREDQTSPPAPGL.....GNHSYCRNPDEDPGPGWCYV 60Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mmc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 357 | 100.0 | 263 | 4 | 000318 |
| 2 | 357 | 100.0 | 263 | 4 | 096FE7 |
| 3 | 155 | 43.4 | 562 | 6 | 08SQ23 |
| 4 | 146 | 40.9 | 516 | 4 | 09BU99 |
| 5 | 146 | 40.9 | 653 | 11 | 08VCS4 |
| 6 | 136 | 38.1 | 559 | 11 | 091VP2 |
| 7 | 131 | 36.7 | 154 | 4 | 096S88 |
| 8 | 130.5 | 36.6 | 717 | 13 | P70006 |
| 9 | 128 | 35.9 | 716 | 11 | F70521 |
| 10 | 127 | 35.6 | 157 | 6 | 09TV48 |
| 11 | 127 | 35.6 | 313 | 13 | 09PU78 |
| 12 | 126 | 35.6 | 395 | 4 | 09BZM1 |
| 13 | 126 | 35.3 | 716 | 11 | 091XG8 |
| 14 | 122.5 | 34.3 | 716 | 13 | 091691 |
| 15 | 122 | 34.2 | 704 | 13 | 090865 |
| 16 | 121 | 33.9 | 810 | 4 | Q15146 |

| | | | | | | | |
|----|-------|------|------|----|--------|--------|--------------|
| 17 | 120 | 33.6 | 812 | 11 | 09R0M3 | 09R0M3 | rattus norv |
| 18 | 118 | 33.1 | 728 | 6 | 09BHO9 | 09BHO9 | felis silve |
| 19 | 118 | 33.1 | 806 | 6 | 018783 | 018783 | macropus eu |
| 20 | 117.5 | 32.9 | 399 | 4 | 096GL8 | 096GL8 | homo sapien |
| 21 | 117.5 | 32.9 | 420 | 4 | 09BTP9 | 09BTP9 | homo sapien |
| 22 | 113 | 31.7 | 452 | 13 | 090Y90 | 090Y90 | xenopus lae |
| 23 | 112.5 | 31.5 | 334 | 5 | 046507 | 046507 | papio hamad |
| 24 | 111 | 31.1 | 385 | 5 | 025101 | 025101 | herdmania m |
| 25 | 109 | 30.5 | 202 | 13 | 090675 | 090675 | gallus gall |
| 26 | 109 | 30.5 | 567 | 4 | 013208 | 013208 | homo sapien |
| 27 | 109 | 30.5 | 685 | 5 | 024488 | 024488 | drosofila |
| 28 | 108.5 | 30.4 | 615 | 6 | 097507 | 097507 | sus scrofa |
| 29 | 108.5 | 30.4 | 812 | 11 | 091M05 | 091M05 | mus musculu |
| 30 | 108 | 30.3 | 115 | 13 | 042341 | 042341 | gallus gall |
| 31 | 107.5 | 30.1 | 1145 | 5 | 09BKL8 | 09BKL8 | apiysia cal |
| 32 | 106.5 | 29.8 | 420 | 13 | 090504 | 090504 | eptarecus |
| 33 | 106 | 29.7 | 359 | 6 | 08WMR1 | 08WMR1 | canis famli |
| 34 | 106 | 29.7 | 378 | 13 | 090WP0 | 090WP0 | tracheus s |
| 35 | 105 | 29.4 | 710 | 13 | 091402 | 091402 | xenopus. he |
| 36 | 105 | 29.4 | 726 | 13 | 090978 | 090978 | gallus gall |
| 37 | 104 | 29.1 | 560 | 4 | 014520 | 014520 | homo sapien |
| 38 | 103 | 28.9 | 608 | 13 | 09PTW7 | 09PTW7 | struthio ca |
| 39 | 102.5 | 28.7 | 597 | 11 | 035727 | 035727 | mus musculu |
| 40 | 102 | 28.6 | 385 | 13 | 090WS2 | 090WS2 | elaphe sp. |
| 41 | 95.5 | 26.8 | 113 | 4 | 09UTRS | 09UTRS | homo sapien |
| 42 | 94 | 26.3 | 709 | 13 | 090ZN6 | 090ZN6 | brachydantio |
| 43 | 93.5 | 26.2 | 594 | 5 | P91823 | P91823 | caenorhabdi |
| 44 | 93 | 26.1 | 214 | 6 | 09XT70 | 09XT70 | oryctolagus |
| 45 | 92.5 | 25.9 | 648 | 4 | Q9H1V4 | Q9H1V4 | homo sapien |

ALIGNMENTS

RESULT 1

AC 000318 PRELIMINARY; PRT; 263 AA.
ID 000318
DR 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE WUSC:DJ515N1.2 protein.
GN WUSC:DJ515N1.2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Du Z., Scheet P., Harper M., PAC clone RP3-515N1."
RT "The sequence of H. sapiens clone RP3-515N1."
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Waterston R.
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC002073; AAB54054.1; -
DR HSSP; P00749; IKDU.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE; FALSE_NEG.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SMO0130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
SQ SEQUENCE 263 AA; 28248 MW; 197C3EEBE54A242 CRC64;

Query Match 100.0%; Score 357; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 2,4e-34;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CFWDNGHLTYREDQTSPPAPGLRLNWLDAQSLGASAPVSGAGNHSYCRNPDEDPGPGWCYV 60
|||||

DB 25 CFWDNGHLYREDQTSAPGLRCLNMLDQSGLASAPVSGAGNHSYCRNPDEDPGRCYV 84

RESULT 2

Q96FE7 ID 096FE7 PRELIMINARY; PRT; 263 AA.

AC 096FE7; 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Unknown (protein for MG:17330).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxId=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=BRIN;

RA Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL; BC011049; AAH11049.1; -.

DR InterPro; IPR000001; Kringle.

DR Pfam; PF00051; Kringle; 1.

DR ProDom; PD000395; Kringle; 1.

DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.

DR PROSITE; PS00070; KRINGLE_2; 1.

SQ SEQUENCE 263 AA; 28234 MW; 197C3EBE88F242 CRC64;

Query Match 100.0%; Score 357; DB 4; Length 263;

Best Local Similarity 100.0%; Pred. No. 2,4e-34;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFWDNGHLYREDQTSAPGLRCLNMLDQSGLASAPVSGAGNHSYCRNPDEDPGRCYV 60

DB 25 CFWDNGHLYREDQTSAPGLRCLNMLDQSGLASAPVSGAGNHSYCRNPDEDPGRCYV 84

RESULT 3

Q8SQ23 ID 08SQ23 PRELIMINARY; PRT; 562 AA.

AC 08SQ23; 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE T-plasminogen activator.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI_TaxId=9823;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=ENMEL ORGAN;

RA Ding Y., Xue J., Bartlett J.D.,

RT "T-plasminogen activator in tooth tissues."

RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL; AF364605; AAM00297.1; -.

SQ SEQUENCE 562 AA; 63668 MW; F9E6B4C77CB101B8 CRC64;

Query Match 43.4%; Score 155; DB 6; Length 562;

Best Local Similarity 43.5%; Pred. No. 4,1e-10;

Matches 30; Conservative 9; Mismatches 18; Indels 12; Gaps 3;

QY 1 CFWDNGHLYREDQTSAPGLRCLNMLDQSGLASAPVSGAGNHSYCRNPDEDPGRCYV 51

DB 127 CYEDQISYRGCTWSTAESGAECTNW--NTSGLASMYNGRRPDAYKLGIGNHNYCRNPDR 184

QY 52 DPGRCYV 60

DB 185 DSK-PWCYI 192

RESULT 4

Q9BU99 ID 09BU99 PRELIMINARY; PRT; 516 AA.

AC 09BU99; 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Similar to plasminogen activator, tissue.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxId=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=SKIN;

RA Strausberg R.;

RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.

CC EMBL; BC002795; AAH02795.1; -.

DR HSSP; P00750; IASH.

DR MEROPS; S01.232; -.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR000561; EGF-like.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR001254; Ser_protease_Try.

DR Pfam; PF00008; EGF; 1.

DR Pfam; PF00051; kringle; 2.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000395; Kringle; 2.

DR SMART; SM00181; EGF; 1.

DR SMART; SM00001; EGF-like; 1.

DR SMART; SM00130; KR_2.

DR SMART; SM00020; Tryp_Spc; 1.

DR PROSITE; PS00022; EGF_1; UNKNOWN_1.

DR PROSITE; PS01186; EGF_2; 1.

DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_2.

DR PROSITE; PS00070; KRINGLE_2; 2.

DR PROSITE; PS00240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW EGF-like domain; glycoprotein; Hydrolyase; Serine protease.

SQ SEQUENCE 516 AA; 57370 MW; BAB31901FDC96800 CRC64;

Query Match 40.9%; Score 146; DB 4; Length 516;

Best Local Similarity 43.5%; Pred. No. 4,4e-09;

Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGHLYREDQTSAPGLRCLNMLDQSGLASAPVSGAGNHSYCRNPDEDPGRCYV 51

DB 81 CYEDQISYRGCTWSTAESGAECTNW--NSSALAKRYSGRRPDARLGLIGNHNYCRNPDR 138

QY 52 DPGRCYV 60

DB 139 DSK-PWCYV 146

RESULT 5

Q8VCS4 ID 08VCS4 PRELIMINARY; PRT; 653 AA.

AC 08VCS4; 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Hypothetical 70.6 kDa protein.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxId=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=LIVER;

RA Strausberg R.;

DB 139 DSK-PWCYV 146

RESULT 5

Q8VCS4 ID 08VCS4 PRELIMINARY; PRT; 653 AA.

AC 08VCS4; 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Hypothetical 70.6 kDa protein.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxId=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=LIVER;

RA Strausberg R.;

Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC019376; AA119376.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001438; EGF_1.
DR InterPro; IPR000083; Fibronectin.
DR InterPro; IPR000562; FN Type_II.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR01254; Ser_protease_Try.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00010; EGFLOOD.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; FN_Type_II; 1.
DR SMART; SM00095; FN_Type_II; 1.
DR SMART; SM00181; EGF_2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_spec; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS01253; FIBRONECTIN_1; UNKNOWN_1.
DR PROSITE; PS00023; FIBRONECTIN_2; UNKNOWN_1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;

Query Match 40.9%; Score 146; DB 11; Length 653;
Best Local Similarity 47.1%; Pred. No. 5.6e-09;
Matches 32; Conservative 7; Mismatches 19; Indels 10; Gaps 3;
QY 1 CPWDNGHLTREQOTSPAPGLRCLNM-----LDAQGLASAPVSGNGNSYGRNPDE 52
DB 283 CFLGNGTERGVASTAASGLSCLANNSDLXYQLHVDV-VAAAVLLGLGPHAYCNPDKD 341
QY 53 PRGPCYV 60
DB 342 ER-PWCYV 348

RESULT 6
Q91VP2 PRELIMINARY; PRT; 559 AA.
AC Q91VP2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to plasminogen activator, tissue.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BREAST TUMOR;
RA Strusberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC011256; AA11256.1; -.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000083; Fibronctn.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.

DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS01253; FIBRONECTIN_1; UNKNOWN_1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 559 AA; 63122 MW; 8CCEB2BD94514D9 CRC64;

Query Match 38.1%; Score 136; DB 11; Length 559;
Best Local Similarity 40.6%; Pred. No. 7.3e-08;
Matches 28; Conservative 8; Mismatches 21; Indels 12; Gaps 3;
QY 1 CPWDNGHLTREQOTSPAPGLRCLNMIDAQGLASAPV-----GAGNHSYGRNPDE 51
DB 124 CFEEOGITYRGWTSTAESGAECLNM--NSSVLSLKPYARRPNAIKLGHNHYCRNPDR 181
QY 52 PRGPCYV 60
DB 182 DLK-PWCYV 189

RESULT 7
Q96SE8 PRELIMINARY; PRT; 154 AA.
ID Q96SE8;
AC Q96SE8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Urokinase-type plasminogen activator amino-terminal fragment.
DE Urokinase-type plasminogen activator amino-terminal fragment.
GN ATF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Fu J., Bai X., Ruan C.;
RT "Cloning and expression of the amino-terminal fragment of human urokinase-type plasminogen activator."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Bai X., Fu J., Wang W., Xi X., Ruan C.;
RT "Overexpression of the amino-terminal fragment of human urokinase-type plasminogen activator in breast cancer cells results in decreased tumor invasion, growth and angiogenesis."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY029537; AA138734.1; -.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Kinase.
SQ SEQUENCE 154 AA; 17305 MW; A3CCF2FCFF505572 CRC64;

Query Match 36.7%; Score 131; DB 4; Length 154;
Best Local Similarity 40.3%; Pred. No. 7.7e-08;
Matches 27; Conservative 7; Mismatches 25; Indels 8; Gaps 2;

OY 1 CFWNGHLYREDQTSAPAGRLCLNWLDA-----QSGLASAPVSGAGNHSYCRNPDEDP 53
 DB 70 CYEGNGHFRGASTDTMGRCPLPWNASATVLOQTYHARSBLQGLGKKNYCRNPD-NR 128
 OY 54 RQPMCYV 60
 DB 129 RQPMCYV 135

RESULT 8

P70006 PRELIMINARY; PRT; 717 AA.
 AC P70006;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hepatocyte growth factor-like protein precursor.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 OC Xenopodinae; Xenopus.
 NC NCB1_TaxID=8355;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RX MEDLINE=96404125; PubMed=8808403;
 RA Abberger F., Schmidt G., Richter K.;
 RT "The Xenopus homologue of hepatocyte growth factor-like protein is
 RT specifically expressed in the presumptive neural plate during
 RT gastrulation.";
 RL Mech. Dev. 54:23-37(1996).
 DR EMBL; Y08734; CAA6989.1; -.
 DR HSSP; P00747; ICEA.
 DR MEROPS; S01.977; -.

DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR003609; Pan_app.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00051; Kringle; 4.
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00018; KRINGLE.
 DR PRODOM; PD000395; Kringle; 4.
 DR SMART; SM00130; KR; 4.
 DR SMART; SM00473; PAN_AP; 1.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_1; 3.
 DR PROSITE; PS50070; KRINGLE_2; 4.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 FT HYDROLASE; Serine protease; Signal.
 FT SIGNAL 1 28
 FT CHAIN 29 717 HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN.
 FT SEQUENCE 717 AA; 82017 MW; 6F877AA432C8DD54 CRC64;

Query Match 36.6%; Score 130.5; DB 13; Length 717;
 Best Local Similarity 38.7%; Pred. No. 4.2e-07;
 Matches 24; Conservative 7; Mismatches 28; Indels 3; Gaps 1;

OY 1 CFWNGHLYREDQTSAPAGRLCLNWLDAQSGT---ASAPVSGAGNHSYCRNPDEDPRGPW 57
 DB 378 CYHNGELIYRGVSKTKRKIGWCRWEKQNDLELSLAQPIYVPLENYCRNPDRSHGPW 437
 OY 58 CY 59
 DB 438 CY 439

RESULT 9

P70521 PRELIMINARY; PRT; 716 AA.
 AC P70521;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Macrophage stimulating protein precursor.
 GN MSP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCB1_TaxID=10116;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=97011126; PubMed=8858136;
 RA Ohshtro K., Iwama A., Matsuno K., Ezaki T., Sakamoto O., Hamaguchi I.,
 RA Takasu N., Suda T.;
 RT "Molecular cloning of Rat Macrophage-stimulating protein and its
 RT involvement in the Male Reproductive System."
 RL Biochem. Biophys. Res. Commun. 227:273-280(1996).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL; X95096; CAA64473.1; -.
 DR HSSP; P00747; IKN.

DR MEROPS; S01.975; -.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR003609; Pan_app.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00051; Kringle; 4.
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRODOM; PD000395; Kringle; 4.
 DR SMART; SM00130; KR; 4.
 DR SMART; SM00473; PAN_AP; 1.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_1; 4.
 DR PROSITE; PS50070; KRINGLE_2; 4.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 KM HYDROLASE; Serine protease; Signal.
 FT SIGNAL 1 31
 FT SEQUENCE 716 AA; 80733 MW; 06B7DF3EF56D921F CRC64;

Query Match 35.9%; Score 128; DB 11; Length 716;
 Best Local Similarity 41.9%; Pred. No. 8.3e-07;
 Matches 26; Conservative 3; Mismatches 29; Indels 4; Gaps 2;

OY 1 CFWNGHLYREDQTSAPAGRLCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPRGPW 57
 DB 110 CIMDNGASYRGTVARTADGLPCQWSRRFPNDHKYTPRKNGL--EENFCRNPDGPPRGPW 168
 OY 58 CY 59
 DB 169 CY 170

RESULT 10
 O9TVAB PRELIMINARY; PRT; 157 AA.

DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Urokinase plasminogen activator (Fragment).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.
 NC NCB1_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SKELETAL MUSCLE;
 RX MEDLINE=21071388; PubMed=11204721;
 RA Balcerzak D., Querengueser L., Dixon W.T., Baracos V.E.;

"Coordinate expression of matrix-degrading proteinases and their RT activators and inhibitors in bovine skeletal muscle.";
RT J. Anim. Sci. 79:94-107(2001).
DR EMBL: AF144761; AAD30301.1; -.
DR HSSP: P00749; IURK.
DR MEROPS; S01.231; -.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Kinase.
FT NON_TER 1 157 157
SQ SEQUENCE 157 AA; 17858 MW; A768D6C72C1FBFB7 CRC64;

Query Match 35.6%; Score 127; DB 6; Length 157;
Best Local Similarity 41.2%; Pred. No. 2.3e-07;
Matches 28; Conservative 6; Mismatches 24; Indels 10; Gaps 3;

QY 1 CFWDNGHLVREDQTSPPAGRLCLNWLDAQSGI-----ASAPVSGAGNHSYCRNPDE 52
DB 38 CYGNGHSYRKARNDLSGRCLAW-DSPTVLLKMYHAHSDAQLQLGKHNVCRNPD-N 95
QY 53 PRGPWCYV 60
DB 96 QRRPCYV 103

RESULT 11
ID Q9PU78 PRELIMINARY; PRT; 313 AA.
AC Q9PU78;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hepatocyte growth factor-like protein (Fragment).
OS Crocodylus niloticus (Nile crocodile) (African crocodile).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylia; Crocodylinae; Crocodylus.
OX NCBI_Taxid=8501;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=20022983; PubMed=10555283;
RA Hughes S., Zelus D., Mouchroud D.;
RT "Warm-blooded isochore structure in nile crocodile and turtle.";
RL Mol. Biol. Evol. 16:1521-1527(1999).
DR EMBL; AJ011396; CAB56422.1; -.
DR HSSP; P00747; IHPJ.
DR MEROPS; S01.977; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
KW Hydrolyase; Serine protease.
FT NON_TER 1 313 313
SQ SEQUENCE 313 AA; 34793 MW; 8E08470495BBSAN2 CRC64;

Query Match 35.6%; Score 127; DB 13; Length 313;
Best Local Similarity 36.5%; Pred. No. 4.7e-07;

Matches 23; Conservative 8; Mismatches 28; Indels 4; Gaps 2;

QY 1 CFWDNGHLVREDQTSPPAGRLCLNWLDAQSGIAS-APVSGAGNHSYCRNPDEPRGP 56
DB 30 CYGNGELVGRHTSKTRGVTCQKWSQSPHYVPSPTTHRAHLDENYCNPNDNSGP 89
QY 57 WCY 59
DB 90 WCY 92

RESULT 12
ID Q9BZM1 PRELIMINARY; PRT; 395 AA.
AC Q9BZM1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Neonatal thrombolytic agent alpha-form (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dou D.;
RT "A brain-type plasminogen activator.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.
DR EMBL; AF260825; AAK11956.1; -.
DR HSSP; P00750; 1PK2.
DR MEROPS; S01.232; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000083; Fibrinctn.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS01283; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_DOM; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolyase; Serine protease.
FT NON_TER 395 395
SQ SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match 35.6%; Score 127; DB 4; Length 395;
Best Local Similarity 39.1%; Pred. No. 5.9e-07;
Matches 27; Conservative 7; Mismatches 23; Indels 12; Gaps 3;

QY 1 CFWDNGHLVREDQTSPPAGRLCLNWLDAQSGIASAPVSGAGNHSYCRNPDE 51
DB 48 CYGNGSAVRGHTSHLSGASCLPNNSMILLIGKYTAQN--PSAALALGKHNVCRNPDG 105
QY 52 DRRGPWCYV 60
DB 106 DAK-PWCHV 113

RESULT 13
ID Q91XG8 PRELIMINARY; PRT; 716 AA.
AC Q91XG8;

DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hepatocyte growth factor-like.
GN HGF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Strubeberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010551; AAH10551.1; -.
DR MGD; MGI:96080; Hgf1.
DR InterPro; IPR00001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR ProDom; PD000395; Kringle; 4.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
KM Hydroxylase; Serine protease.
SQ SEQUENCE 716 AA, 80693 MW, 12474C48A7D4B46D CRC64;

Query Match 35.3%; Score 126; DB 11; Length 716;
Best Local Similarity 41.9%; Pred. No. 1.4e-06;
Matches 26; Conservative 3; Mismatches 29; Indels 4; Gaps 2;

QY 1 CFWNGHLYREDQTSAPGLRCLNWL---LDAQSGLASAPVSGAGNHSYCRNPDEDPGRGPW 57
DB 110 CIMGNGVSYGRVARTAGGLPCQAWSRFPNDHXYTPTPKGL-ENFCRNPDPGDPGRGPW 168
QY 58 CY 59
DB 169 CY 170

RESULT 14
Q91691 PRELIMINARY; PRT; 716 AA.
AC Q91691;

DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Growth factor liverline.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxId=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Ruiz i Altaba A.; They C.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U57455; AAB52574.1; -.
DR HSSP; P00747; ICEA.
DR InterPro; IPR00001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp_Spc; 1.

DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
KM Hydroxylase; Serine protease.
SQ SEQUENCE 716 AA, 81971 MW, 508376A0E4398798 CRC64;

Query Match 34.3%; Score 122.5; DB 13; Length 716;
Best Local Similarity 37.5%; Pred. No. 3.7e-06;
Matches 22; Conservative 8; Mismatches 29; Indels 3; Gaps 1;

QY 1 CFWNGHLYREDQTSAPGLRCLNWLDAQSGL---ASAPVSGAGNHSYCRNPDEDPGRGPW 57
DB 377 CYHNGELYSGRVSKTRKIKCRWEKRNDELSDQYLVPLBENYCRNPDRSHGPW 436
QY 58 CY 59
DB 437 CY 438

RESULT 15
Q90865 PRELIMINARY; PRT; 704 AA.
AC Q90865;

DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hepatocyte growth factor-like/macrophage stimulating protein.
GN HGF1/MSF.

OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxId=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA MEDLINE=96029010; PubMed=7554499;

RT "Expression of HGF/SF, HGF1/MSF and c-met suggests new functions during early chick development.";
RL Dev. Genet. 17:90-101(1995).
RC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE TRYPSIN FAMILY.
CC EMBL; X84043; CAA58862.1; -.
DR HSSP; P00747; ICEA.

DR MEROPS; S01.977; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
KM Hydroxylase; Serine protease.
SQ SEQUENCE 704 AA, 79341 MW, CAB0D8CC41367C37 CRC64;

Query Match 34.2%; Score 122; DB 13; Length 704;
Best Local Similarity 37.3%; Pred. No. 4.2e-06;
Matches 25; Conservative 6; Mismatches 22; Indels 14; Gaps 2;

QY 1 CFWNGHLYREDQTSAPGLRCLNWLDAQSGLASAP-----VSGAGNHSYCRNPDEDPGRGPW 52
DB 108 CIVANGTSYGRTRDITRGLRCQHW-----QATPHDHRFLPLSLRNGLEENYCRNPDRD 161

Mon Apr 7 10:22:32 2003

us-10-057-951-2_copy_25_84.rpt

Page 7

OY 53 PRGPNCT 59
| | | | |
Db 152 KRGPACT 158

Search completed: April 7, 2003, 09:21:36
JOB time : 17.0915 Secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:24 ; Search time 11.2266 Seconds
(without alignments)
712.151 Million cell updates/sec

Title: US-10-057-951-2_COPY_25_84

Perfect score: 357
Sequence: 1 CFMDNGHLVREDQTSAPGL.....GNHSYCNDEDEPRGWCYV 60

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

1: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 357 | 100.0 | 263 | 20 | AAW05219 |
| 2 | 357 | 100.0 | 263 | 20 | AAW05219 |
| 3 | 357 | 100.0 | 263 | 22 | AAW93748 |
| 4 | 357 | 100.0 | 263 | 22 | AAE00300 |
| 5 | 357 | 100.0 | 263 | 23 | AAW06149 |
| 6 | 351 | 98.3 | 263 | 21 | AAW43237 |
| 7 | 343 | 96.1 | 286 | 20 | AAW05220 |
| 8 | 186 | 52.1 | 56 | 20 | AAW12615 |
| 9 | 185 | 51.8 | 55 | 20 | AAW12397 |
| 10 | 183 | 51.3 | 39 | 19 | AAW72641 |

| | | | | | | |
|----|-----|------|-----|----|----------|---------------------|
| 11 | 183 | 51.3 | 39 | 19 | AAW72640 | Nervous glia cell |
| 12 | 149 | 41.7 | 527 | 19 | AAW54154 | t-PA mutant (N142S) |
| 13 | 149 | 41.7 | 527 | 19 | AAW54157 | t-PA mutant (N142S) |
| 14 | 148 | 41.5 | 390 | 9 | AAW82587 | Modified tissue pl |
| 15 | 148 | 41.5 | 483 | 16 | AAW08899 | Human tissue PA va |
| 16 | 148 | 41.5 | 483 | 16 | AAW08877 | Human tissue PA va |
| 17 | 148 | 41.5 | 483 | 16 | AAW08788 | Human tissue PA va |
| 18 | 148 | 41.5 | 483 | 16 | AAW08799 | Human tissue PA va |
| 19 | 148 | 41.5 | 483 | 16 | AAW08800 | Human tissue PA va |
| 20 | 148 | 41.5 | 483 | 16 | AAW08811 | Human tissue PA va |
| 21 | 148 | 41.5 | 483 | 16 | AAW08822 | Human tissue PA va |
| 22 | 148 | 41.5 | 483 | 16 | AAW08833 | Human tissue PA va |
| 23 | 148 | 41.5 | 483 | 16 | AAW08844 | Human tissue PA va |
| 24 | 148 | 41.5 | 483 | 16 | AAW08855 | Human tissue PA va |
| 25 | 148 | 41.5 | 483 | 16 | AAW08866 | Human tissue PA va |
| 26 | 148 | 41.5 | 483 | 16 | AAW08877 | Human tissue PA va |
| 27 | 148 | 41.5 | 483 | 16 | AAW08888 | Human tissue PA va |
| 28 | 148 | 41.5 | 483 | 16 | AAW08899 | Human tissue PA va |
| 29 | 148 | 41.5 | 527 | 13 | AAW21598 | Human tPA variant |
| 30 | 148 | 41.5 | 527 | 14 | AAW44812 | Human tPA variant |
| 31 | 148 | 41.5 | 527 | 14 | AAW44816 | Human tPA variant |
| 32 | 148 | 41.5 | 527 | 14 | AAW44817 | Human tPA variant |
| 33 | 148 | 41.5 | 527 | 16 | AAW70874 | Human t-PA variant |
| 34 | 148 | 41.5 | 527 | 16 | AAW70875 | Human t-PA variant |
| 35 | 148 | 41.5 | 527 | 16 | AAW70876 | Human t-PA variant |
| 36 | 148 | 41.5 | 527 | 16 | AAW70898 | Human t-PA variant |
| 37 | 148 | 41.5 | 527 | 16 | AAW70899 | Human t-PA variant |
| 38 | 148 | 41.5 | 527 | 16 | AAW70900 | Human t-PA variant |
| 39 | 148 | 41.5 | 527 | 16 | AAW70901 | Human t-PA variant |
| 40 | 148 | 41.5 | 527 | 16 | AAW70902 | Human t-PA variant |
| 41 | 148 | 41.5 | 527 | 16 | AAW70903 | Human t-PA variant |
| 42 | 148 | 41.5 | 527 | 16 | AAW70904 | Human t-PA variant |
| 43 | 148 | 41.5 | 527 | 16 | AAW70905 | Human t-PA variant |
| 44 | 148 | 41.5 | 527 | 16 | AAW70906 | Human t-PA variant |
| 45 | 148 | 41.5 | 527 | 16 | AAW70907 | Human t-PA variant |

ALIGNMENTS

| | | |
|----------|-------------|---|
| RESULT 1 | AAW05219 | |
| ID | AAW05219 | standard; Protein; 263 AA. |
| AC | AAW05219; | |
| XX | | |
| DT | 17-JUN-1999 | (first entry) |
| DE | | Kringel1 protein sequence. |
| XX | | |
| KW | | Kringel1; cancer; inflammation; autoimmunity; rheumatoid arthritis; AIDS; |
| KW | | CNS inflammation; cerebellar degeneration; Alzheimer's disease; asthma; |
| KW | | Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis; |
| KW | | neurological abnormality; ischemia reperfusion injury; ischaemic injury; |
| KW | | cardiovascular disease; kidney disease; liver disease; aplastic anaemia; |
| KW | | myocardial infarction; hypotension; hypertension; allergy; infection; |
| KW | | myelodysplastic syndrome; haematologic abnormality; diagnosis; therapy; |
| KW | | male pattern baldness. |
| XX | | |
| OS | | Homo sapiens. |
| XX | | |
| PN | | WO911788-A1. |
| XX | | |
| PD | | 11-MAR-1999. |
| XX | | |
| PF | | 02-SEP-1998; 98WO-US18270. |
| XX | | |
| PR | | 01-SEP-1998; 98US-0144889. |
| XX | | |
| PR | | 02-SEP-1997; 97US-0056032. |
| XX | | |
| PA | | (SMIK) SMITHKLINE BEECHAM CORP. |
| XX | | |
| PI | | Albone EF, Kikly KK; |

```
XX MPI: 1999-214707/18.
DR N-PSDB; AAX28354.
XX
XX New kringle1 polypeptides and polynucleotides
XX
XX Claim 1, Page 31-32; 42pp; English.
XX
CC This sequence is a Kringle1 polypeptide of the invention.
CC The kringle1 polypeptides (I) are used to screen for agonists and
CC antagonists. Agonists are used to treat subjects in need of enhanced
CC activity or expression of (I). Antagonists are used to treat subjects
CC having need to inhibit the activity or expression of (I). The methods can
CC be used to treat conditions such as cancer, inflammation, autoimmunity,
CC allergy, asthma, rheumatoid arthritis, CNS inflammation, cerebellar
CC degeneration, Alzheimer's disease, Parkinson's disease, multiple
CC sclerosis, amyotrophic lateral sclerosis, head injury damage and other
CC neurological abnormalities, ischaemia reperfusion injury, cardiovascular
CC disease, kidney disease, liver disease, ischaemic injury, myocardial
CC infarction, hypertension, AIDS, myelodysplastic syndromes
CC and other haematologic abnormalities, aplastic anaemia, male pattern
CC baldness, and bacterial, fungal, protozoan and viral infections. The
CC kringle1 polypeptides may also be used to generate antibodies.
CC Determining the presence or absence of mutations in, and analysing for
CC the presence or absence of expression of, kringle1 polynucleotides can be
CC used to diagnose a disease or susceptibility to a disease related to
CC expression or activity of kringle1 proteins. The polynucleotides may also
CC be used for chromosome identification, and mapping.
XX
XX Sequence 263 AA;
XX
XX Query Match 100.0%; Score 357; DB 20; Length 263;
XX Best Local Similarity 100.0%; Pred. No. 1,9e-34;
XX Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CFWDNGHL YREDQTS PAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDDPRGPGWCYV 60
XX |||||
XX Db 25 CFWDNGHL YREDQTS PAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDDPRGPGWCYV 84
XX
XX RESULT 2
XX AAM87769
XX ID AAM87769 standard; Protein; 263 AA.
XX
XX AAM87769;
XX
XX 29-MAR-1999 (first entry)
XX
XX Human tissue plasminogen activator-like protease t-PALP.
XX
XX Tissue plasminogen activator-like protease; t-PALP; human;
XX circulatory system-related disorder; blood clotting; stroke;
XX thrombosis; peripheral arterial occlusion; pulmonary embolism;
XX myocardiothrombosis; diagnosis; therapy.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..21
XX FT /label= Sig_peptide
XX FT 22..263
XX FT /label= Mat_protein
XX FT 25..84
XX FT /note= "kringle domain"
XX FT 85..263
XX FT /note= "protease domain"
XX FT 22..31
XX FT /note= "epitope-bearing region"
XX FT 35..44
XX FT /note= "epitope-bearing region"
XX FT 71..81
XX FT /note= "epitope-bearing region"
XX Peptide 91..107
```

```
FT /note= "epitope-bearing region"
FT 119..128
FT /note= "epitope-bearing region"
FT 138..147
FT /note= "epitope-bearing region"
FT 155..167
FT /note= "epitope-bearing region"
FT 193..203
FT /note= "epitope-bearing region"
FT 206..215
FT /note= "epitope-bearing region"
FT 227..237
FT /note= "epitope-bearing region"
FT 243..252
FT /note= "epitope-bearing region"
XX
XX WO9854199-A1.
XX
XX 03-DEC-1998.
XX
XX 27-MAY-1998; 98WO-US10728.
XX
XX 28-MAY-1997; 97US-0048000.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ebner R, Moore PA, Ruben SM;
XX
XX MPI: 1999-070207/06.
XX DR N-PSDB; AAV99636.
XX
XX New tissue plasminogen activator-like protease - useful in the
XX diagnosis and treatment of circulatory system-related disorders
XX
XX Claim 1; Page 56-57; 76pp; English.
XX
XX This is the amino acid sequence of tissue plasminogen activator-like
XX protease (t-PALP), a novel member of the serine protease family
XX that shares sequence homology to human tissue plasminogen activator
XX (see AAM87770). The t-PALP sequence was deduced from a cDNA clone
XX (see AAV99636) derived from activated monocytes. The 2.5 kb t-PALP
XX message has also been detected in heart, brain, lung, placenta,
XX liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate,
XX testis, ovary, small intestine, colon and peripheral blood
XX leukocytes. Isolated nucleic acids encoding amino acids -21 to
XX 242, -20 to 242, 1-242, 4-63 (kringle domain) and 64-242 (protease
XX domain) of t-PALP, or encoding epitope-bearing portions of t-PALP,
XX are also claimed, as are recombinant vectors, host cells, and
XX methods for producing t-PALP polypeptides. t-PALP may be used to
XX detect and treat disorders related to the circulatory system, and
XX to identify agonists and antagonists of t-PALP activity. The
XX homology between t-PALP and tPA indicates that t-PALP may be
XX involved in the regulation of normal and abnormal clotting
XX in e.g. stroke, deep-vein thrombosis, peripheral arterial
XX occlusion, pulmonary embolism and myocardiothrombosis.
XX
XX Sequence 263 AA;
XX
XX Query Match 100.0%; Score 357; DB 20; Length 263;
XX Best Local Similarity 100.0%; Pred. No. 1,9e-34;
XX Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CFWDNGHL YREDQTS PAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDDPRGPGWCYV 60
XX |||||
XX Db 25 CFWDNGHL YREDQTS PAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDDPRGPGWCYV 84
XX
XX RESULT 3
XX AAM93748
XX ID AAM93748 standard; Protein; 263 AA.
XX
XX AAM93748;
XX
```

| | | | |
|--------------------------|---|---------------|--|
| DT | 06-NOV-2001 | (first entry) | |
| DE | Human polypeptide, SEQ ID NO: 3727. | | |
| XX | | | |
| XX | Human, full length cDNA; cDNA synthesis; oligo-capping. | | |
| XX | | | |
| XX | Homo sapiens. | | |
| OS | | | |
| XX | EP1130094-A2. | | |
| XX | | | |
| XX | 05-SEP-2001. | | |
| PD | | | |
| XX | | | |
| XX | 07-JUL-2000; 2000EP-0114089. | | |
| PF | | | |
| XX | 08-JUL-1999; 99JP-0194486. | | |
| PR | 11-JAN-2000; 2000JP-0118774. | | |
| PR | 02-MAY-2000; 2000JP-0183765. | | |
| XX | | | |
| XX | (HELI-) HELIX RES INST. | | |
| XX | | | |
| PI | Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y; | | |
| PI | Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H; | | |
| XX | WPI; 2001-524255/58. | | |
| DR | N-PSDB; AAK94700. | | |
| XX | | | |
| PT | 830 Primers useful for synthesizing full length cDNA clones and their | | |
| PT | use in genetic manipulation - | | |
| XX | | | |
| PS | Claim 8; SEQ ID NO 3727; 1380PP + sequence listing; English. | | |
| XX | | | |
| CC | The invention relates to primers for synthesizing full length cDNA | | |
| CC | clones. 830 cDNA molecules encoding a human protein have been | | |
| CC | isolated and nucleotide sequences of 5' - and 3' -ends of the cDNA | | |
| CC | molecules have been determined. Primers for synthesizing the full length | | |
| CC | cDNA are useful for clarifying the function of the protein encoded by | | |
| CC | the cDNA. The full length clones were obtained by construction of full | | |
| CC | length enriched cDNA libraries that were synthesised by the oligo-capping | | |
| CC | method. The primers enable the production of the full length cDNA easily | | |
| CC | without any special methods. The present invention is a polypeptide | | |
| CC | encoded by a full length human cDNA of the invention. | | |
| CC | Note: The sequence data for this patent did not form part of the printed | | |
| CC | specification, but was obtained in CD-ROM format directly from EPO. | | |
| XX | | | |
| XX | Sequence 263 AA; | | |
| Query Match | 100.0%; Score 357; DB 22; Length 263; | | |
| Best Local Similarity | 100.0%; Pred. No. 1.9e-34; | | |
| Matches 60; Conservative | 100; Mismatches 0; Indels 0; Gaps 0; | | |
| 0y | | | |
| 1 | CPWDGNGHYREDOQTSPAPGLRCLNMLWDAQSGLASAPVSGAGNHSYCRNDDPRGPGWCV 60 | | |
| 25 | CPWDGNGHYREDOQTSPAPGLRCLNMLWDAQSGLASAPVSGAGNHSYCRNDDPRGPGWCV 84 | | |
| Db | | | |
| RESULT 4 | | | |
| AAE00300 | | | |
| ID | AAE00300 standard; Protein; 263 AA. | | |
| XX | | | |
| AC | AAE00300; | | |
| XX | | | |
| DT | 13-JUN-2001 (first entry) | | |
| XX | | | |
| DE | Human tissue-plasminogen activator-like protease (t-PAI). | | |
| XX | | | |
| XX | Human; tissue-plasminogen activator-like protease; t-PAI; | | |
| XX | therapy; vascular disease; stroke; deep vein thrombosis; keloid; asthma; | | |
| XX | arterial calcification; blood coagulation disorder; cerebroprotective; | | |
| XX | autoimmune system disorder; human immunodeficiency syndrome; cytostatic; | | |
| XX | rheumatoid arthritis; graft-versus-host disease; thyroiditis; cardiac; | | |
| XX | insulin dependent diabetes; inflammatory eye disease; allergy; antiviral; | | |
| XX | cardiovascular disease; heart disease; arrhythmia; myocardial ischaemia; | | |
| XX | hyperproliferative disorder; hypertrophic scar; neurological disease; | | |

| | |
|----|--|
| KW | Creutzfeldt-Jakob syndrome; neurodegenerative disorder; antiinflammatory |
| KW | Alzheimer's disease; Parkinson's disease; immunosuppressive; antitungal; |
| KW | infectious disease; drug screening; gene therapy; neuroprotective; |
| KW | cancer; ophthalmological; antibacterial; vulnerary. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| FT | Key |
| FT | Binding-site |
| FT | 1..165 |
| FT | /note= "Binds to FLAG polypeptide to form |
| FT | t-PALP-FLAG fusion protein" |
| FT | 1..21 |
| FT | /label= Signal_peptide |
| FT | 4..63 |
| FT | /label= Kringle_domain |
| FT | 12..21 |
| FT | /note= "Conserved region" |
| FT | 22..263 |
| FT | /note= "Human mature tissue-plasminogen activator-like |
| FT | protease (t-PALP); Binds to FLAG polypeptide to form |
| FT | t-PALP-FLAG fusion protein" |
| FT | 22..38 |
| FT | /note= "Conserved region" |
| FT | 22..31 |
| FT | /note= "Epitope-bearing portion" |
| FT | 35..44 |
| FT | /note= "Epitope-bearing portion" |
| FT | 39..49 |
| FT | /note= "Conserved region" |
| FT | 50..62 |
| FT | /note= "Conserved region" |
| FT | 63..84 |
| FT | /note= "Conserved region" |
| FT | 64..242 |
| FT | /label= Protease_domain |
| FT | 71..81 |
| FT | /note= "Epitope-bearing portion" |
| FT | 85..97 |
| FT | /note= "Conserved region" |
| FT | 91..107 |
| FT | /note= "Epitope-bearing portion" |
| FT | 100..118 |
| FT | /note= "Conserved region" |
| FT | 119..128 |
| FT | /note= "Epitope-bearing portion" |
| FT | 119..127 |
| FT | /note= "Conserved region" |
| FT | 128..143 |
| FT | /note= "Conserved region" |
| FT | 138..147 |
| FT | /note= "Epitope-bearing portion" |
| FT | 146..163 |
| FT | /note= "Conserved region" |
| FT | 155..167 |
| FT | /note= "Epitope-bearing portion" |
| FT | 164..180 |
| FT | /note= "Conserved region" |
| FT | 186..200 |
| FT | /note= "Conserved region" |
| FT | 193..203 |
| FT | /note= "Epitope-bearing portion" |
| FT | 201..220 |
| FT | /note= "Conserved region" |
| FT | 206..215 |
| FT | /note= "Epitope-bearing portion" |
| FT | 221..236 |
| FT | /note= "Conserved region" |
| FT | 227..237 |
| FT | /note= "Epitope-bearing portion" |
| FT | 237..248 |
| FT | /note= "Conserved region" |
| FT | 243..252 |
| FT | /note= "Epitope-bearing portion" |

FT Region 249..263
 /note= "Conserved region"
 XX
 XX WO200125252-A1.
 XX
 XX 12-APR-2001.
 XX
 XX 03-OCT-2000; 2000WO-US27239.
 PF
 XX 04-OCT-1999; 99US-0411977.
 PR
 XX (HUMA-) HUMAN GENOME SCT INC.
 PA
 XX Moore PA, Ruben SM, Edner R;
 PI WPI; 2001-235402/24.
 XX N-PSDB; AAD03460.
 DR
 XX
 XX New (gene encoding and antibody immunospecific for a)
 PT tissue- plasminogen activator-like protease, useful for the diagnosis
 PT and treatment of (cardio)vascular diseases, hyperproliferative
 PT disorders, immune system disorders and cancers -
 PT
 XX Claim 17; Fig 1; 323pp; English.
 XX
 XX The present amino acid sequence is HMS1842 clone human
 CC tissue- plasminogen activator-like protease (t-PALP). The t-PALP
 CC sequence and their (ant)agonists are useful for the diagnosis and
 CC treatment of vascular diseases e.g. stroke, deep vein thrombosis and
 CC arterial occlusion, blood coagulation disorders, (auto)immune system
 CC disorders e.g. human immunodeficiency syndrome, rheumatoid arthritis,
 CC graft-versus-host disease, thyroiditis, insulin dependent diabetes and
 CC inflammatory eye disease, allergic reactions e.g. asthma, cardiovascular
 CC diseases e.g. heart disease, arrhythmia and myocardial ischaemia,
 CC hyperproliferative disorders, cancers, hypertrophic scars and keloids,
 CC neurological diseases e.g. Creutzfeldt-Jakob syndrome, neurodegenerative
 CC disorders e.g. Alzheimer's disease and Parkinson's disease and infectious
 CC disease e.g. viral, bacterial and fungal infections. The t-PALP sequences
 CC are also useful for drug screening. The t-PALP nucleotides are useful as
 CC chromosome markers and are involved in gene therapy.
 XX
 XX Sequence 263 AA;
 SQ
 Query Match 100.0%; Score 357; DB 22; Length 263;
 Best Local Similarity 100.0%; Pred. No. 1.9e-34;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGRCWCYV 60
 DB 25 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGRCWCYV 84
 RESULT 5
 AAU86149
 ID AAU86149 standard; Protein; 263 AA.
 XX
 XX AAU86149;
 XX
 XX 15-JUL-2002 (first entry)
 DT
 XX
 XX Human PRO264 polypeptide.
 DE
 XX
 XX Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;
 KW leukaemia; neuronal disorder; stromal disorder; blastocoeleic disorder;
 KW inflammatory disorder; immune disorder; angiogenic disorder;
 KW cytostatic; neuroprotective.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO200153486-A1.
 PN
 XX 26-JUL-2001.
 PD
 XX

PF 11-FEB-2000; 2000WO-US03565.
 XX
 XX 08-MAR-1999; 99WO-US05028.
 PR 11-MAR-1999; 99US-123972P.
 PR 11-MAY-1999; 99US-133459P.
 PR 02-JUN-1999; 99WO-US12252.
 PR 22-JUN-1999; 99US-140650P.
 PR 22-JUN-1999; 99US-140653P.
 PR 20-JUL-1999; 99US-144758P.
 PR 26-JUL-1999; 99US-146588P.
 PR 28-JUL-1999; 99US-146222P.
 PR 17-AUG-1999; 99US-149395P.
 PR 31-AUG-1999; 99US-151689P.
 PR 01-SEP-1999; 99WO-US20111.
 PR 15-SEP-1999; 99WO-US21099.
 PR 30-NOV-1999; 99WO-US28313.
 PR 01-DEC-1999; 99WO-US28304.
 PR 01-DEC-1999; 99WO-US28634.
 PR 05-JAN-2000; 2000WO-US00219.
 XX
 XX (GETH) GENENTECH INC.
 PA
 XX Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AJ, Hillan KJ;
 PI Masters SA, Pan Y, Pitti RM, Roy MA, Smith V, Stone DW;
 PI Watanabe CK, Wood WI;
 DR WPI; 2002-205567/26.
 XX N-PSDB; ABK40275.
 DR
 XX
 XX Thirty five nucleic acids encoding PRO polypeptides, useful for
 PT treating benign or malignant tumours, leukaemias and lymphoid
 PT malignancies, inflammatory, angiogenic and immunologic disorders -
 PT
 XX Claim 61; Fig 44; 302pp; English.
 PS
 XX The present invention relates to the isolation of novel human PRO
 CC polypeptides and the polynucleotide sequences encoding them. The
 CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are
 CC useful for treating benign or malignant tumours (e.g. renal, kidney,
 CC bladder, breast, etc), leukaemias and lymphoid malignancies, other
 CC disorders such as neuronal, glial, astrocytic, hypothalamic, glandular,
 CC macrophagal, stromal and blastocoeleic disorders, inflammatory, immune
 CC and angiogenic disorders. The polynucleotide sequences are also
 CC useful in gene therapy. AAU86128-AAU86162 represent the human PRO
 CC polypeptides of the invention.
 XX
 XX Sequence 263 AA;
 SQ
 Query Match 100.0%; Score 357; DB 23; Length 263;
 Best Local Similarity 100.0%; Pred. No. 1.9e-34;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGRCWCYV 60
 DB 25 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGRCWCYV 84
 RESULT 6
 AAB43237
 ID AAB43237 standard; Protein; 263 AA.
 XX
 XX AAB43237;
 XX
 XX 08-FEB-2001 (first entry)
 DT
 XX
 XX Human ORFX ORF3001 polypeptide sequence SEQ ID NO:6002.
 DE
 XX
 XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antiporiatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopahic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW

KW antihaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
OS Homo sapiens.
XX
PM WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US08621.
XX
PR 31-MAR-1999; 99US-0127607.
XX
PR 02-APR-1999; 99US-0127636.
XX
PR 05-APR-1999; 99US-0127728.
XX
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
PI WPI: 2000-602362/57.
XX
DR N-PSDB; AAC77446.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
PS Claim 11; Page 5181-5182; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatocytic; vulnary;
CC antiproliferative; antiparkinsonian; nootropic; neuroprotective;
CC osteoprotic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;
CC antichryd; and antihaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 263 AA;
Query Match 98.3%; Score 351; DB 21; Length 263;
Best Local Similarity 98.3%; Pred. No. 9.9e-34;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CPWDGHHYREDQTSFAPGLKCLNWLDAOSGLASAPVSGAGNSYCRNPDEDPGRPCVY 60
DB 25 CPWDGHHYREDQTSFAPGLKCLNWLDAOSGLASAPVSGAGNSYCRNPDEDPGRPCVY 84

RESULT 7
AAV05220
ID AAV05220 standard; Protein: 286 AA.
XX
AC AAV05220;

XX 17-UN-1999 (first entry)
XX
XX Kringle1 protein sequence.
DE
XX Kringle1; cancer; inflammation; autoimmunity; rheumatoid arthritis; AIDS;
XX CNS inflammation; cerebellar degeneration; Alzheimer's disease; asthma;
XX Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
XX neurological abnormality; ischaemia reperfusion injury; ischaemic injury;
XX cardiovascular disease; kidney disease; liver disease; aplastic anaemia;
XX myocardial infarction; hypotension; hypertension; allergy; infection;
XX myelodysplastic syndrome; haematologic abnormality; diagnosis; therapy;
XX male pattern baldness.
XX
OS Homo sapiens.
XX
PM WO9911788-A1.
XX
PD 11-MAR-1999.
XX
PF 02-SEP-1998; 98WO-US18270.
XX
PR 01-SEP-1998; 98US-0144869.
XX
PR 02-SEP-1997; 97US-0056032.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Albane EF, Kirkly KK;
XX
PI WPI: 1999-214707/18.
XX
DR N-PSDB; AAX28355.
XX
XX New kringle1 polypeptides and polynucleotides
PT
XX
PS Claim 14; Page 33; 42pp; English.
XX
XX This sequence is a Kringle1 polypeptide of the invention.
CC The kringle1 polypeptides (I) are used to screen for agonists and
CC antagonists. Agonists are used to treat subjects in need of enhanced
CC activity or expression of (I). Antagonists are used to treat subjects
CC having need to inhibit the activity or expression of (I). The methods can
CC be used to treat conditions such as cancer, inflammation, autoimmunity,
CC allergy, asthma, rheumatoid arthritis, CNS inflammation, cerebellar
CC degeneration, Alzheimer's disease, Parkinson's disease, multiple
CC sclerosis, amyotrophic lateral sclerosis, head injury damage and other
CC neurological abnormalities, ischaemia reperfusion injury, cardiovascular
CC disease, kidney disease, liver disease, ischaemic injury, myocardial
CC infarction, hypotension, hypertension, AIDS, myelodysplastic syndromes
CC and other haematologic abnormalities, aplastic anaemia, male pattern
CC baldness, and bacterial, fungal, protozoan and viral infections. The
CC kringle1 polypeptides may also be used to generate antibodies.
CC Determining the presence or absence of mutations in, and analysing for
CC the presence or absence of expression of, kringle1 polynucleotides can be
CC used to diagnose a disease or susceptibility to a disease related to
CC expression or activity of kringle1 proteins. The polynucleotides may also
CC be used for chromosome identification, and mapping.
XX
SQ Sequence 286 AA;
Query Match 96.1%; Score 343; DB 20; Length 286;
Best Local Similarity 96.7%; Pred. No. 9.7e-33;
Matches 58; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CPWDGHHYREDQTSFAPGLKCLNWLDAOSGLASAPVSGAGNSYCRNPDEDPGRPCVY 60
DB 25 CPWDGHHYREDQTSFAPGLKCLNWLDAOSGLASAPVSGAGNSYCRNPDEDPGRPCVY 84

RESULT 8
AAV12615
ID AAV12615 standard; Protein: 56 AA.
XX
AC AAV12615;

```
XX 22-JUN-1999 (first entry)
XX
XX Human 5' EST secreted protein SEQ ID NO: 280 from WO 9906553.
DE
XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
XX forensic; gene therapy; chromosome mapping; signal peptide;
XX upstream regulatory sequence; cytokine activity; cell proliferation;
XX differentiation; haematopoiesis regulation; tissue growth regulation;
XX reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
XX thrombolytic; anti-inflammatory; tumour inhibition; antitumour.
XX
XX Homo sapiens.
OS
XX
XX WO9906553-A2.
XX
XX 11-FEB-1999.
XX
XX 31-JUL-1998; 98WO-IB01237.
XX
XX 01-AUG-1997; 97US-0905051.
XX
XX (GEST ) GENSET.
XX
XX Duclert A, Dumas Milne Edwards J, Lacroix B;
XX
XX WPI; 1999-153783/13.
XX
XX N-PSDB; AAX41473.
XX
XX New nucleic acids encoding human secreted proteins - obtained from
XX PT cDNA libraries derived from umbilical cord, lymph ganglia,
XX lymphocytes and placental tissue
XX
XX Claim 34; Page 376; 411pp; English.
XX
XX AAX41379 to AAX41526 represent 5' expressed sequence tags (ESTs) for
XX CC human secreted proteins, and encode the proteins given in AAY12521 to
XX CC AAY12668, respectively. The proteins given represent the signal peptide
XX CC and an N-terminal fragment of a secreted protein. The nucleic acid
XX CC sequences can be used for producing secreted human gene products. They
XX CC can also be used to develop products for diagnosis and therapy. The
XX CC proteins obtained may have cytokine activity, cell
XX CC proliferation/differentiation activity, haematopoiesis regulating
XX CC activity, tissue growth regulating activity, reproductive hormone
XX CC regulatory activity, chemotactic/chemokinetic activity, haemostatic and
XX CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX CC activity, tumour inhibition activity or other activities. The products
XX CC can be used in forensic, gene therapy and chromosome mapping procedures.
XX CC The sequences can also be used for obtaining corresponding promoter
XX CC sequences. The nucleic acids encoding the signal peptide can be used for
XX CC directing extracellular secretion of a polypeptide or the insertion of a
XX CC polypeptide into a membrane, or importing a polypeptide into a cell.
XX
XX Sequence 56 AA;
SQ
Query Match 52.1%; Score 186; DB 20; Length 56;
Best Local Similarity 96.9%; Pred. No. 8.3e-15;
Matches 31; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CFWDNGHLYREDQTSPPAGLRCLNWLDAQSG 32
DB 25 CFWDNGHLYREDQTSPPAGLRCLNWLDAQNG 56
RESULT 9
AAY12397
ID AAY12397 standard; Protein; 55 AA.
XX
XX AAY13397;
XX
XX 17-JUN-1999 (first entry)
XX
XX Human 5' EST secreted protein SEQ ID NO:428.
```

```
XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
XX forensic; gene therapy; chromosome mapping; signal peptide;
XX upstream regulatory sequence; cytokine activity; cell proliferation;
XX differentiation; haematopoiesis regulation; tissue growth regulation;
XX reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
XX thrombolytic; anti-inflammatory; tumour inhibition.
XX
XX Homo sapiens.
OS
XX
XX WO9906548-A2.
XX
XX 11-FEB-1999.
XX
XX 31-JUL-1998; 98WO-IB01222.
XX
XX 01-AUG-1997; 97US-0905135.
XX
XX (GEST ) GENSET.
XX
XX Duclert A, Dumas Milne Edwards J, Lacroix B;
XX
XX WPI; 1999-153778/13.
XX
XX N-PSDB; AAX41230.
XX
XX New nucleic acids encoding human secreted proteins - obtained from
XX PT cDNA libraries prepared from e.g. liver, ovary, brain, prostate,
XX kidney, lung, umbilical cord, placenta and colon tissue
XX
XX Claim 27; Page 744; 824pp; English.
XX
XX AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
XX CC human secreted proteins, and encode the proteins given in AAY12261 to
XX CC AAY12514, respectively. The proteins given represent the signal peptide
XX CC and an N-terminal fragment of a secreted protein. The nucleic acid
XX CC sequences can be used for producing secreted human gene products. They
XX CC can also be used to develop products for diagnosis and therapy. The
XX CC proteins obtained may have cytokine activity, cell
XX CC proliferation/differentiation activity, haematopoiesis regulating
XX CC activity, tissue growth regulating activity, reproductive hormone
XX CC regulatory activity, chemotactic/chemokinetic activity, haemostatic and
XX CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX CC activity, tumour inhibition activity or other activities. The products
XX CC can be used in forensic, gene therapy and chromosome mapping procedures.
XX CC The sequences can also be used for obtaining corresponding promoter
XX CC sequences. The nucleic acids encoding the signal peptide can be used for
XX CC directing extracellular secretion of a polypeptide or the insertion of a
XX CC polypeptide into a membrane, or importing a polypeptide into a cell.
XX
XX Sequence 55 AA;
SQ
Query Match 51.8%; Score 185; DB 20; Length 55;
Best Local Similarity 100.0%; Pred. No. 1.1e-14;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CFWDNGHLYREDQTSPPAGLRCLNWLDAQSG 31
DB 25 CFWDNGHLYREDQTSPPAGLRCLNWLDAQSG 55
RESULT 10
AAM72641
ID AAM72641 standard; peptide; 39 AA.
XX
XX AAM72641;
XX
XX 05-JAN-1999 (first entry)
XX
XX Nervous glia cell growth factor N-terminal peptide #2.
XX
XX Nervous glia cell growth factor; human; urine; secretion promoter;
XX choline acetyltransferase activity enhancer; nervous disease.
XX
```

```

OS Homo.sapiens.
XX Key Location/Qualifiers
FH Misc-difference 25 /note= "unspecified"
FT Misc-difference 29 /note= "unspecified"
FT Misc-difference 29 /note= "unspecified"
XX JPI0265498-A.
XX 06-OCT-1998.
XX 24-MAR-1997; 97JP-0090305.
XX 24-MAR-1997; 97JP-0090305.
XX 24-MAR-1997; 97JP-0090305.
XX (NICH-) JAPAN CHEM RES CO LTD.
XX WPI; 1998-589719/50.
XX Nervous glia cell growth factor derived from human urine - used for
XX treatment of nervous diseases
XX Claim 3; Fig 7; 14pp; Japanese.
XX The present invention describes nervous glia cell growth factor, which
XX is purified from human urine by ultrafiltration, salting-out by ammonium
XX sulphate, gel filtration, ion exchange chromatography and reversed phase
XX chromatography, and has a M.W. of 29 kDa by sodium dodecyl sulphate-
XX polyacrylamide gel electrophoresis. Also described are: (1) a secretion
XX promoter for the nerve growth factor of glia cell consisting of the
XX above growth factor, an enhancer for choline acetyltransferase activity
XX of neuron consisting of the above growth factor; and (2) DNA encoding
XX nervous glia growth factor containing a DNA sequence coding the amino
XX acid sequence shown by the two 39 amino acid sequences as given in Tyr
XX AAW72640 and AAW72641, which are identical, except one starts with Tyr
XX and the other with Ser (i.e. they are from different DNA transcripts).
XX The glia cell growth factor can be prepared in a large amount and the
XX factor can be used for the treatment of nervous diseases.
XX Sequence 39 AA;
XX
Query Match 51.3%; Score 183; DB 19; Length 39;
Best Local Similarity 94.3%; Pred. No. 1.3e-14;
Matches 33; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 FWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAP 36
DB 5 FWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAP 39

```

RESULT 11

```

AAW72640
ID AAW72640 standard; peptide, 39 AA.
XX
XX AAW72640;
XX
XX 05-JAN-1999 (first entry)
XX
XX Nervous glia cell growth factor N-terminal peptide #1.
XX
XX Nervous glia cell growth factor; human; urine; secretion promoter;
XX choline acetyltransferase activity enhancer; nervous disease.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Misc-difference 25 /note= "unspecified"
FT Misc-difference 29 /note= "unspecified"
XX JPI0265498-A.

```

```

XX
XX 06-OCT-1998.
XX 24-MAR-1997; 97JP-0090305.
XX 24-MAR-1997; 97JP-0090305.
XX 24-MAR-1997; 97JP-0090305.
XX (NICH-) JAPAN CHEM RES CO LTD.
XX WPI; 1998-589719/50.
XX Nervous glia cell growth factor derived from human urine - used for
XX treatment of nervous diseases
XX Claim 2; Fig 6; 14pp; Japanese.
XX The present invention describes nervous glia cell growth factor, which
XX is purified from human urine by ultrafiltration, salting-out by ammonium
XX sulphate, gel filtration, ion exchange chromatography and reversed phase
XX chromatography, and has a M.W. of 29 kDa by sodium dodecyl sulphate-
XX polyacrylamide gel electrophoresis. Also described are: (1) a secretion
XX promoter for the nerve growth factor of glia cell consisting of the
XX above growth factor, an enhancer for choline acetyltransferase activity
XX of neuron consisting of the above growth factor; and (2) DNA encoding
XX nervous glia growth factor containing a DNA sequence coding the amino
XX acid sequence shown by the two 39 amino acid sequences as given in
XX AAW72640 and AAW72641, which are identical, except one starts with Tyr
XX and the other with Ser (i.e. they are from different DNA transcripts).
XX The glia cell growth factor can be prepared in a large amount and the
XX factor can be used for the treatment of nervous diseases.
XX Sequence 39 AA;
XX
Query Match 51.3%; Score 183; DB 19; Length 39;
Best Local Similarity 94.3%; Pred. No. 1.3e-14;
Matches 33; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 FWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAP 36
DB 5 FWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAP 39

```

RESULT 12

```

AAW54154
ID AAW54154 standard; protein, 527 AA.
XX
XX AAW54154;
XX
XX 20-JUL-1998 (first entry)
XX
XX t-PA mutant (N142S).
XX
XX Amino acid substitution; t-PA; vascular disorder; prevention;
XX fibrin deposition; adhesion formation.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Domain 1..44 /note= "Finger domain"
FT Domain 45..91 /note= "Growth factor domain"
FT Domain 92..173 /note= "Kring1-1 domain"
FT Domain 180..261 /note= "Kring1-2 domain"
FT Domain 264..527 /note= "Serine protease domain"
FT misc_difference 142 /note= "N changed from wt to S in mutant"
XX
XX US5736135-A.

```

| | |
|----|--|
| PD | 07-APR-1998. |
| XX | |
| PR | 13-FEB-1995; 95US-0389615. |
| XX | |
| PR | 11-JUL-1991; 91US-0728456. |
| PR | 26-JAN-1993; 93US-0008940. |
| PR | 01-APR-1994; 94US-0221660. |
| PR | 13-FEB-1995; 95US-0389615. |
| XX | |
| PA | (GETH) GENENTECH INC. |
| XX | |
| P1 | Goeddel DV, Leung DWH, Rice GC; |
| XX | |
| DR | WP1; 1998-239153/21. |
| XX | |
| PT | Mutant tissue plasminogen activator proteins - useful for treating vascular disorders, preventing tissue adhesion(s), etc. |
| XX | |
| PS | Claim 2; Page -: 24pp; English. |
| XX | |
| CC | Mutant tissue plasminogen activator proteins (AAW5147-W54158) are created by single or multiple amino acid substitutions. Compositions containing the t-PA variant are used for treating vascular disorders, for preventing fibrin deposition or for preventing adhesion formation or reformation. Note: This sequence is not given in the specification but was created from the wild type by the indexer. |
| XX | |
| SQ | Sequence 527 AA; |

```

Query Match 149: DB 19: Length 527;
Query Similarity 41.7%; Pred. No. 2, 4e-09;
Best Local Similarity 41.9%; Pred. No. 2, 4e-09;
Matches 31; Conservative 4; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFMDNGHLYREDTSPAPGLRCMLWLDQSLASAPVS-----GAGNRYCRNPDE 51
      : : : : :
Db 92 CYEDQGISRGWSTAESGAECTWN--NSSALAKPKYSGRRPDALRLGKNSHSCRNPDR 149
      : : : : :
QY 52 DPRGPMCYV 60
      : : : : :
Db 150 DSK-PMCYV 157

```

| | |
|-----------|---|
| RESULT 14 | |
| AAP82587 | |
| ID | AAP82587 standard; protein; 390 AA. |
| XX | |
| AC | |
| XX | AAP82587; |
| XX | |
| DT | 03-NOV-1990 (first entry) |
| DE | Modified tissue plasminogen activator lacking F and G region and |
| XX | kinin1e region 2 and Q-96, I-98 and S-119 substd for N, T and W resp. |
| KW | Modified tissue plasminogen activator; tPA; thrombosis; |
| XX | N-glycosylation. |
| FN | JP63230084-A. |
| XX | |
| PD | 26-SEP-1988. |
| XX | |
| PF | 20-MAR-1987; 87JP-0064340. |
| XX | |
| PR | 20-MAR-1987; 87JP-0064339. |
| XX | |
| XX | (EISA) EISA KK. |
| XX | |
| DR | WPI; 1988-311962/44. |
| XX | |
| PT | N-PSDB; AAN82187. |
| XX | |
| XX | Multiple modified tissue plasminogen activator - which lacks F and G |
| XX | regions, is useful in treatment of thrombosis and has improved half life. |
| XX | |
| XX | disclosure:) : 29pp; Japanese. |

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:19:28 ; Search time 4.24116 Seconds
(without alignments)
864.895 Million cell updates/sec

Title: US-10-057-951-2_COPY_25_84

Perfect score: 1 CFWDNGHLYREDQTSAPGL.....GNHSYCRNPDEDRGWCYV 60

Sequence: BLOSUM62

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PC7_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PC7US_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------------|---------------------|
| 1 | 357 | 100.0 | 263 | 9 US-10-102-704-2 | Sequence 2, Appl1 |
| 2 | 357 | 100.0 | 263 | 9 US-10-057-951-2 | Sequence 2, Appl1 |
| 3 | 357 | 100.0 | 263 | 10 US-09-084-491A-2 | Sequence 2, Appl1 |
| 4 | 146 | 40.9 | 527 | 9 US-09-987-457-18 | Sequence 18, Appl1 |
| 5 | 146 | 40.9 | 527 | 9 US-09-987-455-19 | Sequence 19, Appl1 |
| 6 | 146 | 40.9 | 527 | 9 US-09-974-298-145 | Sequence 145, Appl1 |
| 7 | 146 | 40.9 | 527 | 10 US-09-969-271-7 | Sequence 7, Appl1 |
| 8 | 131 | 36.7 | 88 | 10 US-09-880-503-1 | Sequence 9, Appl1 |
| 9 | 131 | 36.7 | 96 | 10 US-09-880-503-9 | Sequence 9, Appl1 |
| 10 | 131 | 36.7 | 135 | 10 US-09-880-503-4 | Sequence 4, Appl1 |
| 11 | 131 | 36.7 | 138 | 9 US-10-237-667-12 | Sequence 12, Appl1 |
| 12 | 131 | 36.7 | 138 | 9 US-10-237-708-12 | Sequence 12, Appl1 |
| 13 | 131 | 36.7 | 138 | 9 US-10-237-866-12 | Sequence 12, Appl1 |
| 14 | 131 | 36.7 | 138 | 9 US-10-237-871-12 | Sequence 12, Appl1 |
| 15 | 131 | 36.7 | 138 | 10 US-09-984-186-12 | Sequence 12, Appl1 |
| 16 | 131 | 36.7 | 143 | 10 US-09-880-503-8 | Sequence 8, Appl1 |
| 17 | 131 | 36.7 | 403 | 10 US-09-880-503-6 | Sequence 6, Appl1 |
| 18 | 131 | 36.7 | 411 | 10 US-09-880-503-3 | Sequence 3, Appl1 |
| 19 | 131 | 36.7 | 431 | 9 US-10-076-421-2 | Sequence 2, Appl1 |

| | | | | | |
|----|-----|------|-----|---------------------|--------------------|
| 20 | 131 | 36.7 | 431 | 10 US-09-264-468B-1 | Sequence 1, Appl1 |
| 21 | 127 | 35.6 | 326 | 9 US-10-057-951-3 | Sequence 3, Appl1 |
| 22 | 127 | 35.6 | 354 | 9 US-09-987-457-10 | Sequence 10, Appl1 |
| 23 | 127 | 35.6 | 354 | 9 US-09-987-455-11 | Sequence 11, Appl1 |
| 24 | 127 | 35.6 | 372 | 9 US-10-102-704-3 | Sequence 3, Appl1 |
| 25 | 127 | 35.6 | 372 | 10 US-09-084-491A-3 | Sequence 3, Appl1 |
| 26 | 127 | 35.6 | 777 | 9 US-09-987-455-8 | Sequence 8, Appl1 |
| 27 | 121 | 33.9 | 79 | 9 US-09-335-325-8 | Sequence 8, Appl1 |
| 28 | 121 | 33.9 | 79 | 10 US-10-131-241-8 | Sequence 8, Appl1 |
| 29 | 121 | 33.9 | 79 | 10 US-09-753-064-2 | Sequence 2, Appl1 |
| 30 | 121 | 33.9 | 79 | 10 US-09-761-120-8 | Sequence 8, Appl1 |
| 31 | 121 | 33.9 | 160 | 9 US-09-335-325-35 | Sequence 35, Appl1 |
| 32 | 121 | 33.9 | 160 | 9 US-10-131-241-35 | Sequence 35, Appl1 |
| 33 | 121 | 33.9 | 160 | 10 US-09-761-120-35 | Sequence 35, Appl1 |
| 34 | 121 | 33.9 | 250 | 9 US-09-335-325-30 | Sequence 30, Appl1 |
| 35 | 121 | 33.9 | 250 | 9 US-10-131-241-30 | Sequence 30, Appl1 |
| 36 | 121 | 33.9 | 250 | 10 US-09-761-120-30 | Sequence 30, Appl1 |
| 37 | 121 | 33.9 | 260 | 9 US-10-131-241-61 | Sequence 61, Appl1 |
| 38 | 121 | 33.9 | 339 | 9 US-09-335-325-3 | Sequence 3, Appl1 |
| 39 | 121 | 33.9 | 339 | 9 US-10-131-241-3 | Sequence 3, Appl1 |
| 40 | 121 | 33.9 | 339 | 10 US-09-788-142-3 | Sequence 3, Appl1 |
| 41 | 121 | 33.9 | 339 | 10 US-09-761-120-3 | Sequence 3, Appl1 |
| 42 | 121 | 33.9 | 352 | 9 US-09-335-325-40 | Sequence 40, Appl1 |
| 43 | 121 | 33.9 | 352 | 9 US-10-131-241-40 | Sequence 40, Appl1 |
| 44 | 121 | 33.9 | 352 | 10 US-09-761-120-40 | Sequence 40, Appl1 |
| 45 | 121 | 33.9 | 368 | 10 US-09-761-120-42 | Sequence 42, Appl1 |

ALIGNMENTS

RESULT 1
US-10-102-704-2
; Sequence 2, Application US/10102704
; Patent No. US20020164768A1
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protein
; FILE REFERENCE: PF378C1
; CURRENT APPLICATION NUMBER: US/10/102,704
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/084,491
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: 60/048,000
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-102-704-2

Query Match 100.0%; Score 357; DB 9; Length 263;
Best Local Similarity 100.0%; Pred. No. 4.9e-32;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFWDNGHLYREDQTSAPGLRCLNMLDAOSGLASAPVGAAGHNSYCRNPDEDRGWCYV 60
DB 25 CFWDNGHLYREDQTSAPGLRCLNMLDAOSGLASAPVGAAGHNSYCRNPDEDRGWCYV 84

RESULT 2
US-10-057-951-2
; Sequence 2, Application US/10057951
; Patent No. US20020177213A1
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: PF378P1
; CURRENT APPLICATION NUMBER: US/10/057,951
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 09/411,977

;; PRIOR FILING DATE: 1999-10-04
;; PRIOR APPLICATION NUMBER: US 09/084,491
;; PRIOR FILING DATE: 1998-05-27
;; PRIOR APPLICATION NUMBER: US 60/048,000
;; PRIOR FILING DATE: 1997-05-28
;; NUMBER OF SEQ ID NOS: 31
;; SOFTWARE: PatentIn Ver. 3.1
;; SEQ ID NO 2
;; LENGTH: 263
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-057-951-2

Query Match 100.0%; Score 357; DB 9; Length 263;
Best Local Similarity 100.0%; Pred. No. 4.9e-32;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFWDNGLYREDQTSPPAGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGPMCYV 60
DB 25 CFWDNGLYREDQTSPPAGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGPMCYV 84

RESULT 3
US-09-084-491A-2

;; Sequence 2, Application US/09084491A
;; Patent No. US2002061576A1
;; GENERAL INFORMATION:
;; APPLICANT: MOORE, PAUL A.
;; APPLICANT: RUBEN, STEVEN M.
;; APPLICANT: BENNER, REINHARD
;; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
;; NUMBER OF SEQUENCES: 16
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
;; STREET: 9410 KEY WEST AVENUE
;; CITY: ROCKVILLE
;; STATE: MD
;; COUNTRY: USA
;; ZIP: 20850
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/084,491A
;; FILING DATE: 27-MAY-1998
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BROOKES, ANDERS A.
;; REGISTRATION NUMBER: 36,373
;; REFERENCE/DOCKET NUMBER: PF378
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (301) 309-8504
;; TELEFAX: (301) 309-8439
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 263 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-09-084-491A-2

Query Match 100.0%; Score 357; DB 10; Length 263;
Best Local Similarity 100.0%; Pred. No. 4.9e-32;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFWDNGLYREDQTSPPAGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGPMCYV 60
DB 25 CFWDNGLYREDQTSPPAGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGPMCYV 84

RESULT 4

US-09-987-457-18

;; Sequence 19, Application US/09987457
;; Publication No. US20030013150A1
;; GENERAL INFORMATION:
;; APPLICANT: Manosroi, Aranya
;; APPLICANT: Manosroi, Jiradej
;; APPLICANT: Tayapiwatana, Chatchai
;; APPLICANT: Goetz, Friedrich
;; APPLICANT: Werner, Rolf-Guenther
;; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
;; FILE REFERENCE: 0652.2180001
;; CURRENT APPLICATION NUMBER: US/09/987,457
;; CURRENT FILING DATE: 2001-11-14
;; PRIOR APPLICATION NUMBER: 60/268,573
;; PRIOR FILING DATE: 2001-02-15
;; PRIOR APPLICATION NUMBER: GB 00 27 782.2
;; PRIOR FILING DATE: 2000-11-14
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 18
;; LENGTH: 527
;; TYPE: PRT
;; ORGANISM: Homo sapiens (LPA)
US-09-987-457-18

Query Match 40.9%; Score 146; DB 9; Length 527;
Best Local Similarity 43.5%; Pred. No. 1.9e-08;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGLYREDQTSPPAGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGPMCYV 51
DB 92 CYEDQGISYRGTWSTASGAECTNM--NSSALAKKYSRRRDALRLGNNHYCRNPDR 149

QY 52 DPGPMCYV 60
DB 150 DSK-PMCYV 157

RESULT 5
US-09-987-455-19

;; Sequence 19, Application US/09987455
;; Publication No. US20030049729A1
;; GENERAL INFORMATION:
;; APPLICANT: Aranya Manosroi
;; APPLICANT: Jiradej Manosroi
;; APPLICANT: Chatchai Tayapiwatana
;; APPLICANT: Friedrich Goetz
;; APPLICANT: Rolf-Guenther Werner
;; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
;; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
;; FILE REFERENCE: 0652.2190001
;; CURRENT APPLICATION NUMBER: US/09/987,455
;; CURRENT FILING DATE: 2001-11-14
;; PRIOR APPLICATION NUMBER: 60/268,574
;; PRIOR FILING DATE: 2001-02-15
;; PRIOR APPLICATION NUMBER: GB 0027779.8
;; PRIOR FILING DATE: 2000-11-14
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 19
;; LENGTH: 527
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-987-455-19

Query Match 40.9%; Score 146; DB 9; Length 527;
Best Local Similarity 43.5%; Pred. No. 1.9e-08;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGLYREDQTSPPAGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGPMCYV 51
DB 92 CYEDQGISYRGTWSTASGAECTNM--NSSALAKKYSRRRDALRLGNNHYCRNPDR 149

```
QY      52  DFRGPMCYV  60
          | : |||||
Db      150 DSK-PMCYV  157
```

RESULT 6

```

US-09-974-298-145
Sequence 145, Application US/05974298
Patent No. US20020156263A1
GENERAL INFORMATION:
APPLICANT: Chen, Huei-Mei
TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
FILE REFERENCE: PA-0037 P
CURRENT APPLICATION NUMBER: US/09/974,298
CURRENT FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/238,331
PRIOR FILING DATE: 2000-05-10
NUMBER OF SEQ ID NOS: 194
SOFTWARE: PERL Program
SEQ ID NO 145
LENGTH: 562
TYPE: PR1
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CD1
US-09-974-298-145

```

| Query Match | Score | DB | Length |
|-------------|-------|----|--------|
| 40.98; | 146; | 9; | 562; |

DQ 1 CFMNGHLTRDQTSAPGLRCLNWLDAQSLASAPV-----GAGHNSYCRPDE 51
| : | | | : | | | | | : | | |
Db 127 CYEDQGISYRGTWSTAESGAECTNM--NSSALAQKPYSGRRDAIRLGLGNHNYCRNPDR 184

| | | | |
|----|-----|-----------|-----|
| QY | 52 | DPRGPMCYV | 60 |
| | | : | |
| Db | 185 | DSK-PMCYV | 192 |

RESULT 7

```

US-09-969-271-7
; Sequence 7, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB))
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-969-271-7

```

| Query Match | Score | DB | Length |
|-------------|-------|----|--------|
| 40.9% | 146 | 10 | 562 |

Dy 1 CFNDNGHLRYEDQDTAPAGLRCLNWLDAQSGLASAPV-----GAGHNSYCRNPDE 51
| : | | | | | | | | | |
Db 127 CYEDDGISYRGTSWAESGAECTNW--NSSALAQPYSGRRPDATRLGLGNHYCNRPR 184

QY 52 DPRGPWCYV 60
| : |||||
Db 185 DSK-PWCYV 192

RESULT 8
US-09-88

```

? Sequence 1, Application US/09880503
? Patent No. US20020131964A1
? GENERAL INFORMATION:
? APPLICANT: CHINESE, Douglas B
? APPLICANT: HIGAZI, Abd Al-Rooof
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
? TITLE OF INVENTION: TISSUE CONTRACTABILITY
? FILE REFERENCE: 9596-331
? CURRENT APPLICATION NUMBER: US/09/880,503
? CURRENT FILING DATE: 2001-06-13
? PRIOR APPLICATION NUMBER: US 60/212,847
? PRIOR FILING DATE: 2000-06-20
? NUMBER OF SEQ ID NOS: 18
? SOFTWARE: Patentin Ver. 2.1
? SEQ ID NO 1
? LENGTH: 86
? TYPE: PRT
? ORGANISM: Homo sapiens
? US-09-880-503-1

```

| | | | | |
|--------------------------|--------|--------------------|-----------|------------|
| Query Match | 36.7%; | Score 131; | DB 10; | Length 88; |
| Best Local Similarity | 40.3%; | Pred. No. 1.3e-07; | | |
| Matches 27; Conservative | 7; | Mismatches 25; | Indels 8; | Gaps 2 |

QY 1 CFMDNDHLYKREDQTSBPGLRCLAMLDL-----QSGLSAPVPSGAGNSYCRNDDSD 54
 Db 3 CYGNGNHFFRKGASLSTDMGRFCLPMSNATYVQOTYHAHRSDALQGLGKSNYCRNPD-NR 61
 QY 54 RCPWCTV 60
 Db 62 RRPWCYV 68

| | | | |
|----|----|---------|----|
| QY | 54 | RGPWCYV | 60 |
| | | | |
| Db | 62 | RRPWCYV | 68 |

RESULT 9

```

US-09-880-503-9
; Sequence 9, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Boof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR APPLICATION NUMBER: 2001-06-13
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 96
; TYPE: prt
; ORGANISM: Homo sapiens
; IS-09-880-503-9

```

| | | | | |
|--------------------------|--------|--------------------|-----------|------------|
| Query Match | 36.7%; | Score 131; | DB 10; | Length 96; |
| Best Local Similarity | 40.3%; | Pred. No. 1.5e-07; | | |
| Matches 27; Conservative | 7; | Mismatches 25; | Indels 8; | Gaps 22; |

QY 1 CFWDNGHLXREDQTSBAFGLRCLTNWLDA-----QSGLASAPVSGAGNHSYCRNPDED 53

Dp 3 CTEGNGHFYRGAKASTDTMGRLPWN SATVLQQTVAHRRSDALQGLGKENVCRNPD-NR 61

| | | | |
|----|----|---------|----|
| Qy | 54 | RGPWCYV | 60 |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| Db | 62 | RRPWCYV | 68 |

RESULT 10

TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
MOLECULE TYPE: linear
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-708-12

Query Match 36.7%; Score 131; DB 9; Length 138;
Best Local Similarity 40.3%; Pred. No. 2.2e-07;
Matches 27; Conservative 7; Mismatches 25; Indels 8; Gaps 2;

QY 1 CFWDNGHLVREDQTSAPARCLNMLDA-----OSGLASAPVSGAGNHSYCRNPDEDP 53
DB 53 CYEGNGHFRGKASTDTWGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPD-NR 111
QY 54 RGPWCYV 60
DB 112 RRPWCYV 118

RESULT 13
US-10-237-866-12
Sequence 12, Application US/10237866
Publication No. US20030036171A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcoia Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,866
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-866-12

Query Match 36.7%; Score 131; DB 9; Length 138;
Best Local Similarity 40.3%; Pred. No. 2.2e-07;
Matches 27; Conservative 7; Mismatches 25; Indels 8; Gaps 2;

QY 1 CFWDNGHLVREDQTSAPARCLNMLDA-----OSGLASAPVSGAGNHSYCRNPDEDP 53
DB 53 CYEGNGHFRGKASTDTWGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPD-NR 111
QY 54 RGPWCYV 60
DB 112 RRPWCYV 118

RESULT 14
US-10-237-871-12
Sequence 12, Application US/10237871
Publication No. US20030036172A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcoia Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,871
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-871-12

Search completed: April 7, 2003, 09:31:28
Job time : 4.2416 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:17:08 ; Search time 51.2682 Seconds
(Without alignments)
754.542 Million cell updates/sec

Title: US-10-057-951-2_COPY_25_84
Perfect score: 357
Sequence: 1 CFMNGHLYREDQTSAPGL.....GNHSYCRNDEDPGRCVCYV 60

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 64473110*residues
Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Pending Patents_AA Main:*

- 1: /cgn2_6/ptodata/1/paa/US06 COMB.pdp.*
- 2: /cgn2_6/ptodata/1/paa/US06 COMB.pdp.*
- 3: /cgn2_6/ptodata/1/paa/US07 COMB.pdp.*
- 4: /cgn2_6/ptodata/1/paa/US08 COMB.pdp.*
- 5: /cgn2_6/ptodata/1/paa/US08 COMB.pdp.*
- 6: /cgn2_6/ptodata/1/paa/US08 COMB.pdp.*
- 7: /cgn2_6/ptodata/1/paa/US08 COMB.pdp.*
- 8: /cgn2_6/ptodata/1/paa/US08 COMB.pdp.*
- 9: /cgn2_6/ptodata/1/paa/US08 COMB.pdp.*
- 10: /cgn2_6/ptodata/1/paa/US08 COMB.pdp.*
- 11: /cgn2_6/ptodata/1/paa/US08 COMB.pdp.*
- 12: /cgn2_6/ptodata/1/paa/US08 COMB.pdp.*
- 13: /cgn2_6/ptodata/1/paa/US08 COMB.pdp.*
- 14: /cgn2_6/ptodata/1/paa/US09 COMB.pdp.*
- 15: /cgn2_6/ptodata/1/paa/US09 COMB.pdp.*
- 16: /cgn2_6/ptodata/1/paa/US09 COMB.pdp.*
- 17: /cgn2_6/ptodata/1/paa/US09 COMB.pdp.*
- 18: /cgn2_6/ptodata/1/paa/US09 COMB.pdp.*
- 19: /cgn2_6/ptodata/1/paa/US09 COMB.pdp.*
- 20: /cgn2_6/ptodata/1/paa/US09 COMB.pdp.*
- 21: /cgn2_6/ptodata/1/paa/US09 COMB.pdp.*
- 22: /cgn2_6/ptodata/1/paa/US09 COMB.pdp.*
- 23: /cgn2_6/ptodata/1/paa/US09 COMB.pdp.*
- 24: /cgn2_6/ptodata/1/paa/US10 COMB.pdp.*
- 25: /cgn2_6/ptodata/1/paa/US10 COMB.pdp.*
- 26: /cgn2_6/ptodata/1/paa/US10 COMB.pdp.*
- 27: /cgn2_6/ptodata/1/paa/US60 COMB.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 357 | 100.0 | 109 | 22 | US-09-834-366-13839 |
| 2 | 357 | 100.0 | 109 | 27 | US-60-197-873-13839 |
| 3 | 357 | 100.0 | 117 | 1 | PCT-US01-14827-8897 |
| 4 | 357 | 100.0 | 126 | 17 | US-09-307-140-979 |
| 5 | 357 | 100.0 | 126 | 22 | US-09-817-076-979 |
| 6 | 357 | 100.0 | 146 | 1 | PCT-US02-05301-237 |

| | | | | | | |
|----|-----|-------|-----|----|----------------------|----------------------|
| 7 | 357 | 100.0 | 146 | 1 | PCT-US02-05301-237 | Sequence 311, App |
| 8 | 357 | 100.0 | 178 | 27 | US-60-213-800-237 | Sequence 237, App |
| 9 | 357 | 100.0 | 250 | 14 | US-09-001-403-64 | Sequence 64, App |
| 10 | 357 | 100.0 | 263 | 1 | PCT-US00-27239-2 | Sequence 2, App |
| 11 | 357 | 100.0 | 263 | 1 | PCT-US02-05301-164 | Sequence 164, App |
| 12 | 357 | 100.0 | 263 | 1 | PCT-US98-10728-2 | Sequence 2, App |
| 13 | 357 | 100.0 | 263 | 14 | US-09-084-491A-2 | Sequence 2, App |
| 14 | 357 | 100.0 | 263 | 14 | US-09-144-889-2 | Sequence 2, App |
| 15 | 357 | 100.0 | 263 | 20 | US-09-611-526-3727 | Sequence 3727, App |
| 16 | 357 | 100.0 | 263 | 21 | US-09-791-537-39562 | Sequence 39562, App |
| 17 | 357 | 100.0 | 263 | 23 | US-09-927-726-44 | Sequence 44, App |
| 18 | 357 | 100.0 | 263 | 24 | US-10-057-951-2 | Sequence 44, App |
| 19 | 357 | 100.0 | 263 | 25 | US-10-102-704-2 | Sequence 2, App |
| 20 | 357 | 100.0 | 263 | 26 | US-10-210-951-44 | Sequence 44, App |
| 21 | 357 | 100.0 | 263 | 26 | US-10-211-858-44 | Sequence 44, App |
| 22 | 357 | 100.0 | 263 | 26 | US-10-211-858-44 | Sequence 44, App |
| 23 | 357 | 100.0 | 263 | 27 | US-60-048-000-2 | Sequence 2, App |
| 24 | 357 | 100.0 | 263 | 27 | US-60-230-435-1580 | Sequence 1580, App |
| 25 | 357 | 100.0 | 263 | 27 | US-60-389-987-244 | Sequence 244, App |
| 26 | 357 | 100.0 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 27 | 357 | 100.0 | 308 | 27 | US-60-207-315-471 | Sequence 471, App |
| 28 | 354 | 99.2 | 295 | 1 | PCT-US01-14827-8900 | Sequence 8900, App |
| 29 | 343 | 96.1 | 286 | 15 | US-09-144-889-4 | Sequence 4, App |
| 30 | 306 | 85.7 | 257 | 1 | PCT-US01-14827-8898 | Sequence 8898, App |
| 31 | 186 | 52.1 | 56 | 13 | US-08-905-051-280 | Sequence 280, App |
| 32 | 186 | 52.1 | 56 | 13 | US-08-905-051-280 | Sequence 2338, App |
| 33 | 185 | 51.8 | 55 | 13 | US-08-905-135-428 | Sequence 428, App |
| 34 | 185 | 51.8 | 55 | 13 | US-08-905-135-428 | Sequence 2126, App |
| 35 | 154 | 43.1 | 566 | 21 | US-09-791-537-140530 | Sequence 140530, App |
| 36 | 149 | 41.7 | 527 | 18 | US-09-422-601-1 | Sequence 1, App |
| 37 | 146 | 40.9 | 82 | 3 | US-07-929-1818-44 | Sequence 44, App |
| 38 | 146 | 40.9 | 399 | 27 | US-60-208-020-156 | Sequence 156, App |
| 39 | 146 | 40.9 | 399 | 27 | US-60-209-043-206 | Sequence 206, App |
| 40 | 146 | 40.9 | 472 | 13 | US-08-984-462-2 | Sequence 2, App |
| 41 | 146 | 40.9 | 472 | 13 | US-08-984-462-2 | Sequence 2, App |
| 42 | 146 | 40.9 | 472 | 14 | US-09-032-861-12 | Sequence 12, App |
| 43 | 146 | 40.9 | 472 | 14 | US-09-032-861-12 | Sequence 12, App |
| 44 | 146 | 40.9 | 472 | 14 | US-09-032-861-12 | Sequence 12, App |
| 45 | 146 | 40.9 | 472 | 14 | US-09-079-343-2 | Sequence 2, App |

ALIGNMENTS

US-09-834-366-13839
Sequence 13839, Application US/09834366
GENERAL INFORMATION:
APPLICANT: Bejani, Stephane
APPLICANT: Tanaka, Hiroaki
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Jobert, Severin
APPLICANT: Giordano, Jean-Yves
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: 81 US2 REG
CURRENT APPLICATION NUMBER: US/09/834,366
CURRENT FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/197, 873
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 52153
SOFTWARE: Patent.pm
SEQ ID NO 13839
LENGTH: 109
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -21...-1
US-09-834-366-13839
Query Match 100.0%; Score 357; DB 22; Length 109;
Best Local Similarity 100.0%; Pred. No. 8,1e-34;
Matches 60; Conservative 0; Mismatches 0; Gaps 0;

Oy 1 CFWDNGHLYREQTS PARGRLCLNWLDAQSLASAPVSGAGNHSCRNDEDPREGWCYV 60
|||
|||
|||
|||
|||
Dd 25 CFWNDGHLTYREDQTS PARGLRCLNLWLDQAQLASAPVSGAGNHSCYNRPDEDPRGPWCYV 84
|||
|||
|||
|||
|||

```

RESULT 2
US-60-197-873-13839
: SEQUENCE 13839, Application US/60197873
: GENERAL INFORMATION:
: APPLICANT: Bejani, Stephanie
: APPLICANT: Tanaka, Hitokaki
: APPLICANT: Dumas Mline Edwards, Jean Baptiste
: APPLICANT: Jobert, Severin
: APPLICANT: Giordano, Jean-Yves
: TITLE OF INVENTION: ESTs and Encoded Human Proteins
: FILE REFERENCE: 81, US1, PRO
: CURRENT APPLICATION NUMBER: US/60/197, 873
: CURRENT FILING DATE: 2000-04-18
: NUMBER OF SEQ ID NOS: 52153
: SOFTWARE: Patent.pgm
: SEQ ID NO: 13839
: LENGTH: 109
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SIGNAL
: LOCATION: 21..1
: US-60-197-873-13839

```

| | | | | |
|--------------------------|---------|--------------------|-----------|-------------|
| Query Match | 100.0%; | Score 357; | DB 27; | Length 109; |
| Best Local Similarity | 100.0%; | Pred. No. 8.1e-34; | | |
| Matches 60; Conservative | 0; | Mismatches 0; | Indels 0; | Gaps 0; |

Dy 1 CFWDNGHLYREQDTS P A G L R C L N W I D A Q S L A S A P V S G A G N H S Y C R N D E D D P R G M C Y V 600

Db 25 CFWDNGHLYREQDTS P A G L R C L N W I D A Q S L A S A P V S G A G N H S Y C R N D E D D P R G M C Y V 840

```

RESULT 3
PCT-US01-14827-8897
? Sequence 8897, Application PC/TUS0114827
? GENERAL INFORMATION:
?   APPLICANT: Hyseq, Inc
?   TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
?   FILE REFERENCE: 21272-104
?   CURRENT APPLICATION NUMBER: PCT/US01/14827
?   CURRENT FILING DATE: 2001-05-16
?   PRIOR APPLICATION NUMBER: 09/577,408
?   PRIOR FILING DATE: 2000-05-18
?   NUMBER OF SEQ ID NOS: 16102
?   SOFTWARE: Custom
?   SEQ ID NO 8897
?   LENGTH: 117
?   TYPE: PRT
?   ORGANISM: Homo sapiens
?   FEATURE:
?     NAME/KEY: DOMAIN
?     LOCATION: (74)..(95)
?   OTHER INFORMATION: KRINGLE DOMAIN SIGNATURE domain identified by EMATRIX.
?   OTHER INFORMATION: accession number PR00018C, p-value=6.447e-17, raw score of 14.30
?   NAME/KEY: DOMAIN
?     LOCATION: (36)..(113)
?   OTHER INFORMATION: Kringle domain identified by Pfam, accession name kringle, E-
?     OTHER INFORMATION: value=1.2e-05, Pfam score of 12.2
?   PCT-US01-14827-8897

```

```

Query Match      100.0% Score 357; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 8.8e-34;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 CFWDNGHLYEDDTPAPGLRCINTLMDAOSGLASAPVSGAGNHSYCRNPEDPGPGCYV 60
      * |||||||||||||||||||||||||||||||||||||||||||||||||||||||

```

Db 36 CFWDNGHLYREDQTSAPAGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPRGPWCYV 95

```

RESULT 4
US-09-307-140-979
1 Sequence 979, Application US/09307140
GENERAL INFORMATION:
APPLICANT: Geating, David P.
APPLICANT: Holtsman, Douglas A.
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: Nucleic Acid Molecules Derived from
TITLE OF INVENTION: Human Aortic Endothelium Library
FILE REFERENCE: MM98-14PA
CURRENT APPLICATION NUMBER: US/09/307.140
CURRENT FILING DATE: 1999-05-07
EARLIER APPLICATION NUMBER: 60/084,565
EARLIER FILING DATE: 1998-05-07
NUMBER OF SEQ ID NOS: 1168
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 979
LENGTH: 126
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(21)
US-09-307-140-979

```

| | | | | |
|-----------------------|---------|--------------------|--------|---------------|
| Query Match | 100.0%; | Score 357; | DB 17; | Length 126; |
| Best Local Similarity | 100.0%; | Pred. No. 9.5e-34; | | |
| Matches | 60; | Conservative | 0; | Mismatches 0; |
| | | | Indels | 0; |
| | | | Gaps | 0; |

QY 1 CFWDNGHLYREQDTSAPGLRCLNWLDAQSLASAPVSGAGNHSYCRNDEDEDPRGPCWY 600
25 CFWDNGHLYREQDTSAPGLRCLNWLDAQSLASAPVSGAGNHSYCRNDEDEDPRGPCWY 840

```

RESULT 5
US-09-817-076-979
: Sequence 979, Application US/09817076
: GENERAL INFORMATION:
: APPLICANT: Geating, David P.
: APPLICANT: Holtzman, Douglas A.
: APPLICANT: Robison, Keith E.
: TITLE OF INVENTION: Nucleic Acid Molecules Derived from e
: TITLE OF INVENTION: Human Aortic Endothelium Library
: FILE REFERENCE: MN98-142A
: CURRENT APPLICATION NUMBER: US/09/817,076
: PRIOR FILING DATE: 2001-03-26
: PRIOR APPLICATION NUMBER: 09/307,140
: PRIOR FILING DATE: 1999-05-07
: PRIOR APPLICATION NUMBER: 60/084,565
: PRIOR FILING DATE: 1998-05-07
: NUMBER OF SEQ ID NOS: 1168
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 979
: LENGTH: 126
: TYPE: prt
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SIGNAL
: LOCATION: (1)...(21)
US-09-817-076-979

```

| | | | | |
|-----------------------|----------------|--------------|----------|------------|
| Query Match | 100.0% | Score 357 | DB 22 | Length 126 |
| Best Local Similarity | 100.0% | Pred. No. 9 | 5e-34 | |
| Matches 60 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 |

| | | | |
|----|---|---|----|
| QY | 1 | CPMDNGHLYREDTSAPGRLCLNMLDQSLAAPPVSGAGNHSYCNPPEDRGWCTV | 60 |
| | | CPMDNGHLYREDTSAPGRLCLNMLDQSLAAPPVSGAGNHSYCNPPEDRGWCTV | 84 |

```
RESULT 6 -
PCT-US02-05301-237
; Sequence 237, Application PC/TUS0205301
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PS736PCT
; CURRENT APPLICATION NUMBER: PCT/US02/05301
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/304,417
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/270,625
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 237
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (133)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (137)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (142)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (144)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US02-05301-237

Query Match          100.0%; Score 357; DB 1; Length 146;
Best Local Similarity 100.0%; Pred. No. 1,1e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 60
Db 25 CFWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 84

RESULT 7
PCT-US02-05301-311
; Sequence 311, Application PC/TUS0205301
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PS736PCT
; CURRENT APPLICATION NUMBER: PCT/US02/05301
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/304,417
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/270,625
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 311
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (133)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (137)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (142)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
PCT-US02-05301-311
```

```
; LOCATION: (144)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US02-05301-311

Query Match          100.0%; Score 357; DB 1; Length 146;
Best Local Similarity 100.0%; Pred. No. 1,1e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 60
Db 25 CFWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 84

RESULT 8
US-60-213-800-237
; Sequence 237, Application US/60213800
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: US8 THEREOF
; FILE REFERENCE: CLO00707
; CURRENT APPLICATION NUMBER: US/60/213,800
; CURRENT FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 657
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 237
; LENGTH: 178
; TYPE: PRT
; ORGANISM: HUMAN
US-60-213-800-237

Query Match          100.0%; Score 357; DB 27; Length 178;
Best Local Similarity 100.0%; Pred. No. 1,4e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 60
Db 25 CFWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 84

RESULT 9
US-09-001-403-64
; Sequence 64, Application US/09001403
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Au-Young, Janice
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yue, Henry
; APPLICANT: Shah, Purni
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY PROTEINS
; NUMBER OF SEQUENCES: 150
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/001,403
; FILING DATE: HERREWITH
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
```

```

; NAME: BILLINGS, LUCY J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0455 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LINDNOT05
; CLONE: 3122252
; US-09-001-403-64

Query Match          100.0%; Score 357; DB 14; Length 250;
Best Local Similarity 100.0%; Pred. No. 2e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPRGPMCYV 60
DB 12 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPRGPMCYV 71

RESULT 10
PCT-US00-27239-2
; Sequence 2, Application PC/TUS0027239
; GENERAL INFORMATION:
; APPLICANT: HUMAN GENOME SCIENCES, INC.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: PF378PCT2
; CURRENT APPLICATION NUMBER: PCT/US00/27239
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 09/411,977
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US00-27239-2

Query Match          100.0%; Score 357; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 2.1e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPRGPMCYV 60
DB 25 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPRGPMCYV 84

RESULT 11
PCT-US02-05301-164
; Sequence 164, Application PC/TUS0205301
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PS736PCT
; CURRENT APPLICATION NUMBER: PCT/US02/05301
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/304,417
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/270,625
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 164
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```

PCT-US02-05301-164

Query Match          100.0%; Score 357; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 2.1e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPRGPMCYV 60
DB 25 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPRGPMCYV 84

RESULT 12
PCT-US98-10728-2
; Sequence 2, Application PC/TUS9810728
; GENERAL INFORMATION:
; APPLICANT: EBNER, REINHARD
; APPLICANT: MOORE, PAUL
; APPLICANT: RUBEN, STEVEN
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/10728
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF379PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US98-10728-2

Query Match          100.0%; Score 357; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 2.1e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPRGPMCYV 60
DB 25 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPRGPMCYV 84

RESULT 13
US-09-084-491A-2
; Sequence 2, Application US/09084491A
; GENERAL INFORMATION:
; APPLICANT: MOORE, PAUL A.
; APPLICANT: RUBEN, STEVEN M.
; APPLICANT: EBNER, REINHARD
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
```

Mon Apr 7 10:22:29 2003

us-10-057-951-2_copy_25_84.ram

Page 5

```
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/084,491A
FILING DATE: 27-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PP378
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8439
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-084-491A-2
```

```
Query Match 100.0%; Score 357; DB 14; Length 263;
Best Local Similarity 100.0%; Pred. No. 2.1e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CFMDNGHLYREDQTSAPGRCINWLDNOSGLASAPVSGAGNHSYCRNPDDPRGPGWCYV 60
DB 25 CFMDNGHLYREDQTSAPGRCINWLDNOSGLASAPVSGAGNHSYCRNPDDPRGPGWCYV 84
```

RESULT 14

```
US-09-144-889-2
Sequence 2, Application US/09144889B
GENERAL INFORMATION:
APPLICANT: Bari P. Albone
APPLICANT: Kristine K. Kikly
TITLE OF INVENTION: KRINGLE1
FILE REFERENCE: GH-70249
CURRENT APPLICATION NUMBER: US/09/144,889B
CURRENT FILING DATE: 1998-09-01
EARLIER APPLICATION NUMBER: 60/056,032
EARLIER FILING DATE: 1997-09-02
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-144-889-2
```

```
Query Match 100.0%; Score 357; DB 15; Length 263;
Best Local Similarity 100.0%; Pred. No. 2.1e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CFMDNGHLYREDQTSAPGRCINWLDNOSGLASAPVSGAGNHSYCRNPDDPRGPGWCYV 60
DB 25 CFMDNGHLYREDQTSAPGRCINWLDNOSGLASAPVSGAGNHSYCRNPDDPRGPGWCYV 84
```

RESULT 15

```
US-09-611-526-3727
Sequence 3727, Application US/09611526
GENERAL INFORMATION:
APPLICANT: OTA, TOSHIO
APPLICANT: NISHIKAWA, TETSUO
APPLICANT: ISOGAI, TAKAO
APPLICANT: HAYASHI, KOJI
```

```
APPLICANT: ISHII, SHIZUKO
APPLICANT: KAWAI, YUKI
APPLICANT: WAKAMATSU, AI
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: NAGAI, KEIICHI
APPLICANT: KOJIMA, SHINICHI
APPLICANT: OTSUKI, TETSUJI
APPLICANT: KOGA, HISASHI
TITLE OF INVENTION: PRIMERS FOR SYNTHESIS OF FULL LENGTH CDNAS
TITLE OF INVENTION: AND THEIR USES
FILE REFERENCE: 08335/0122
CURRENT APPLICATION NUMBER: US/09/611,526
CURRENT FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: JP 1999-194486
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: JP 2000-118774
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183765
PRIOR FILING DATE: 2000-05-02
NUMBER OF SEQ ID NOS: 4484
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3727
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-09-611-526-3727
```

```
Query Match 100.0%; Score 357; DB 20; Length 263;
Best Local Similarity 100.0%; Pred. No. 2.1e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CFMDNGHLYREDQTSAPGRCINWLDNOSGLASAPVSGAGNHSYCRNPDDPRGPGWCYV 60
DB 25 CFMDNGHLYREDQTSAPGRCINWLDNOSGLASAPVSGAGNHSYCRNPDDPRGPGWCYV 84
```

Search completed: April 7, 2003, 09:29:27
Job time : 52.2682 secs

10

LENGTH: 264
TYPE: PRT
ORGANISM: Mus musculus
US-10-144-779-488

Query Match 78.4%; Score 280; DB 6; Length 264;
Best Local Similarity 75.0%; Pred. No. 1.3e-21;
Matches 45; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 1 CFWDNGHLYREDQTSPPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 60
DB 25 CFWDNGHLYREDQTSPPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 84

RESULT 3
US-60-452-680-14406

Sequence 14406, Application US/60452680
GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01450

CURRENT FILING DATE: 2003-03-07

NUMBER OF SEQ ID NOS: 116213

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 14406

LENGTH: 293

TYPE: PRT

ORGANISM: Homo sapiens

US-60-452-680-14406

Query Match 40.9%; Score 146; DB 7; Length 293;
Best Local Similarity 43.5%; Pred. No. 1.6e-07;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGHLYREDQTSPPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 51
DB 127 CYEDGISTYRGWTSTASGAECTNW--NSSALAKPKYSGRRPDAIRLGLGNHNYCRNPDR 184

QY 52 DPRGWCYV 60
DB 185 DSK-PWCYV 192

RESULT 4
US-60-453-135-8958

Sequence 8958, Application US/60453135
GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01456

CURRENT FILING DATE: 2003-03-10

NUMBER OF SEQ ID NOS: 82762

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 8958

LENGTH: 293

TYPE: PRT

ORGANISM: Homo sapiens

US-60-453-135-8958

Query Match 40.9%; Score 146; DB 7; Length 293;
Best Local Similarity 43.5%; Pred. No. 1.6e-07;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGHLYREDQTSPPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 51
DB 127 CYEDGISTYRGWTSTASGAECTNW--NSSALAKPKYSGRRPDAIRLGLGNHNYCRNPDR 184

QY 52 DPRGWCYV 60
DB 185 DSK-PWCYV 192

RESULT 5
US-60-453-050-8958

Sequence 8958, Application US/60453050
GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01457

CURRENT FILING DATE: 2003-03-10

NUMBER OF SEQ ID NOS: 82762

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 8958

LENGTH: 293

TYPE: PRT

ORGANISM: Homo sapiens

US-60-453-050-8958

Query Match 40.9%; Score 146; DB 7; Length 293;
Best Local Similarity 43.5%; Pred. No. 1.6e-07;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGHLYREDQTSPPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 51
DB 127 CYEDGISTYRGWTSTASGAECTNW--NSSALAKPKYSGRRPDAIRLGLGNHNYCRNPDR 184

QY 52 DPRGWCYV 60
DB 185 DSK-PWCYV 192

RESULT 6
US-60-452-680-14405

Sequence 14405, Application US/60452680
GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01450

CURRENT FILING DATE: 2003-03-07

NUMBER OF SEQ ID NOS: 116213

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 14405

LENGTH: 516

TYPE: PRT

ORGANISM: Homo sapiens

US-60-452-680-14405

Query Match 40.9%; Score 146; DB 7; Length 516;
Best Local Similarity 43.5%; Pred. No. 2.6e-07;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGHLYREDQTSPPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 51
DB 81 CYEDGISTYRGWTSTASGAECTNW--NSSALAKPKYSGRRPDAIRLGLGNHNYCRNPDR 138

QY 52 DPRGWCYV 60
DB 139 DSK-PWCYV 146

RESULT 7
US-60-453-135-8957

Sequence 8957, Application US/60453135
GENERAL INFORMATION:

APPLICANT: CARGILL, Michele
APPLICANT: IAKUBOVA, Olga
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
CURRENT APPLICATION NUMBER: CL001456
CURRENT FILING DATE: 2003-03-10
NUMBER OF SEQ ID NOS: 82762
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8957
LENGTH: 516
TYPE: PRT
ORGANISM: Homo sapiens
US-60-453-135-8957

Query Match 40.9%; Score 146; DB 7; Length 516;
Best Local Similarity 43.5%; Pred. No. 2.6e-07;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGHLYREDQTSAPGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 51
DB 81 CYEDGISTYRGWSTAESGAECTNW--NSSALAKRYSGRPPDAIRLGLGNHNYCRNPDR 138

QY 52 DPGPCYV 60
DB 139 DSK-PWCYV 146

RESULT 8
US-60-453-050-8957
Sequence 8957, Application US/60453050
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: CLO01457
CURRENT APPLICATION NUMBER: US/60/453,050
CURRENT FILING DATE: 2003-03-10
NUMBER OF SEQ ID NOS: 82762
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8957
LENGTH: 516
TYPE: PRT
ORGANISM: Homo sapiens
US-60-453-050-8957

Query Match 40.9%; Score 146; DB 7; Length 516;
Best Local Similarity 43.5%; Pred. No. 2.6e-07;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGHLYREDQTSAPGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 51
DB 81 CYEDGISTYRGWSTAESGAECTNW--NSSALAKRYSGRPPDAIRLGLGNHNYCRNPDR 138

QY 52 DPGPCYV 60
DB 139 DSK-PWCYV 146

RESULT 9
US-09-612-314A-51
Sequence 51, Application US/09612314A
GENERAL INFORMATION:
APPLICANT: SMITH, RICHARD ANTHONY GODWIN
APPLICANT: DODD, IAN
TITLE OF INVENTION: MOSSAKOMSKA, DANUTA EWA IRENA
FILE REFERENCE: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
CURRENT APPLICATION NUMBER: 37945-0004
CURRENT FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: US/09/612,314A
PRIOR APPLICATION NUMBER: US 09/214,913

PRIOR FILING DATE: 1999-03-16
PRIOR APPLICATION NUMBER: PCT/EP97/03715
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: GB 96 148 71.3
PRIOR FILING DATE: 1996-07-15
NUMBER OF SEQ ID NOS: 53
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 51
LENGTH: 527
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: tissue plasminogen activator
US-09-612-314A-51

Query Match 40.9%; Score 146; DB 5; Length 527;
Best Local Similarity 43.5%; Pred. No. 2.8e-07;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGHLYREDQTSAPGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 51
DB 92 CYEDGISTYRGWSTAESGAECTNW--NSSALAKRYSGRPPDAIRLGLGNHNYCRNPDR 149

QY 52 DPGPCYV 60
DB 150 DSK-PWCYV 157

RESULT 10
PCT-US02-32263-26
Sequence 26, Application PC/TUS0232263
GENERAL INFORMATION:
APPLICANT: Neose Technologies, Inc.
APPLICANT: Defrees, Shawn
APPLICANT: Zopf, David
APPLICANT: Bayer, Robert
APPLICANT: Bower, Caryn
APPLICANT: Hakes, David
TITLE OF INVENTION: REMODELING AND GLYCOCONJUGATION OF PEPTIDES
FILE REFERENCE: 040853-01-5050WO
CURRENT APPLICATION NUMBER: PCT/US02/32263
CURRENT FILING DATE: 2002-10-31
PRIOR APPLICATION NUMBER: US 60/328,523
PRIOR FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: US 60/344,692
PRIOR FILING DATE: 2001-10-19
PRIOR APPLICATION NUMBER: US 60/334,233
PRIOR FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: US 60/334,301
PRIOR FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: US 60/387,292
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: US 60/391,777
PRIOR FILING DATE: 2002-06-25
PRIOR APPLICATION NUMBER: US 60/396,594
PRIOR FILING DATE: 2002-07-17
PRIOR APPLICATION NUMBER: US 60/404,249
PRIOR FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: US 60/407,527
PRIOR FILING DATE: 2002-08-28
NUMBER OF SEQ ID NOS: 62
SOFTWARE: PatentIn version 3.1
SEQ ID NO 26
LENGTH: 562
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-32263-26

Query Match 40.9%; Score 146; DB 1; Length 562;
Best Local Similarity 43.5%; Pred. No. 2.8e-07;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

```

Oy      1  CFWNDGHLHYREDQTSPPAGRLCLNMLDNOQGLASAPVS-----GAGNHSYCRNPDE 51
         |:::|||||
Db      127  CYEDGSGISYRGTWSTAESGABCTNM--NSSALAQKPYSGRRPDATRLG|GNHNHYCRNPDR 184
         |:::|||||

Oy      52  DPRGWCYV 60
         |:::|||||
Db      185  DSK-PWCYV 192

RESULT 11
US-09-703-695A-4
; Sequence 4, Application US/09703695A
; GENERAL INFORMATION:
; APPLICANT: Xu, Yuan
; TITLE OF INVENTION: REVERSE-PHASE HPLC ASSAY FOR PLASMINOGEN ACTIVATORS
; FILE REFERENCE: P1788R1
; CURRENT APPLICATION NUMBER: US/09/703,695A
; CURRENT FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/163,607
; PRIOR FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 4
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-703-695A-4

Query Match      40.9%; Score 146; DB 5; Length 562;
Best Local Similarity 43.5%; Pred. No. 2.8e-07;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3,

Oy      1  CFWNDGHLHYREDQTSPPAGRLCLNMLDNOQGLASAPVS-----GAGNHSYCRNPDE 51
         |:::|||||
Db      127  CYEDGSGISYRGTWSTAESGABCTNM--NSSALAQKPYSGRRPDATRLG|GNHNHYCRNPDR 184
         |:::|||||

Oy      52  DPRGWCYV 60
         |:::|||||
Db      185  DSK-PWCYV 192

RESULT 12
US-10-287-994-26
; Sequence 26, Application US/10287994
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Defrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Bowe, Caryn
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; TITLE OF INVENTION: REMODELING AND GLYCOCOMMUNIGATION OF PEPTIDES
; FILE REFERENCE: 040853-01-5052-00
; CURRENT APPLICATION NUMBER: US/10/287,994
; CURRENT FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 562

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-994-26

Query Match                               40.9%; Score 146; DB 6; Length 562;
Best Local Similarity 43.5%; Pred. No. 2.8e-07;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY      1 CFMNGHLYREDOTSAPGLRCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 51
       | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      127 CYEDOGISYRGWTWSTAESGAECTNW--NSSALAQKPSGRPPAIRLGI GNHN YCRNPDR 184

QY      52 DPRGPWCYV 60
       | : | | | |
Db      185 DSK-PWCYV 192

RESULT 13
US-60-452-680-14408
; Sequence 14408, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GRUPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14408
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-452-680-14408

Query Match                               40.9%; Score 146; DB 7; Length 562;
Best Local Similarity 43.5%; Pred. No. 2.8e-07;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY      1 CFMNGHLYREDOTSAPGLRCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 51
       | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      127 CYEDOGISYRGWTWSTAESGAECTNW--NSSALAQKPSGRPPAIRLGI GNHN YCRNPDR 184

QY      52 DPRGPWCYV 60
       | : | | | |
Db      185 DSK-PWCYV 192

RESULT 14
US-60-453-135-8960
; Sequence 8960, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKUBOVA, Olga
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8960
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-8960

Query Match                               40.9%; Score 146; DB 7; Length 562;
Best Local Similarity 43.5%; Pred. No. 2.8e-07;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY      1 CFMNGHLYREDOTSAPGLRCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 51

```

```

Db      127  CYPDGGISYRGWSTAESGAECTNW--NSSALAKQPYSGRRPDALRLGAGNNHYCRNPDR 184
QY      52   DPRGPMCTV 60
          | : |||
          185   DSK-PMCTV 192
    
```

```

RESULT 15
US-60-453-050-8960
; Sequence 8960, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001457
; CURRENT FILING DATE: 2003-03-10
; CURRENT APPLICATION NUMBER: US/60/453,050
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8960
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-8960
    
```

```

Query Match          40.9%; Score 146; DB 7; Length 562;
Best Local Similarity 43.5%; Pred. No. 2.8e-07;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;
    
```

```

QY      1   CFMNGHLYREDQSPARGLRCLMWLDQSGIASAPVS-----GAGNHSYCRNPDE 51
          | : |||
          127  CYPDGGISYRGWSTAESGAECTNW--NSSALAKQPYSGRRPDALRLGAGNNHYCRNPDR 184
QY      52   DPRGPMCTV 60
          | : |||
          185   DSK-PMCTV 192
    
```

Search completed: April 7, 2003, 09:30:48
 Job time : 9.98129 secs

10

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:28 ; Search time 5.98753 Seconds
(without alignments)
294.842 Million cell updates/sec

Title: US-10-057-951-2_COPY_25_84

Perfect score: 357
Sequence: 1 CFWDNGHLYREDQTSFAPGL.....GNHSTCRNPDDEPRGWCYV 60

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/Backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1 | 357 | 100.0 | 263 | 4 US-09-411-977-2 | Sequence 2, Appl |
| 2 | 150 | 42.0 | 472 | 2 US-08-811-949-63 | Sequence 63, Appl |
| 3 | 146 | 40.9 | 83 | 2 US-08-811-949-2 | Sequence 2, Appl |
| 4 | 146 | 40.9 | 437 | 2 US-08-811-949-49 | Sequence 49, Appl |
| 5 | 146 | 40.9 | 437 | 2 US-08-811-949-51 | Sequence 51, Appl |
| 6 | 146 | 40.9 | 437 | 2 US-08-811-949-55 | Sequence 55, Appl |
| 7 | 146 | 40.9 | 437 | 2 US-08-811-949-57 | Sequence 57, Appl |
| 8 | 146 | 40.9 | 527 | 1 US-07-609-5108-16 | Sequence 16, Appl |
| 9 | 146 | 40.9 | 527 | 2 US-08-811-949-39 | Sequence 39, Appl |
| 10 | 146 | 40.9 | 527 | 5 PCT-US91-01025A-2 | Sequence 2, Appl |
| 11 | 146 | 40.9 | 527 | 5 5185259-8 | Patent No. 5185259 |
| 12 | 146 | 40.9 | 527 | 6 520913-1 | Patent No. 520913 |
| 13 | 146 | 40.9 | 546 | 6 5200340-6 | Patent No. 5200340 |
| 14 | 146 | 40.9 | 562 | 2 US-08-811-949-43 | Sequence 43, Appl |
| 15 | 146 | 40.9 | 562 | 2 US-08-560-098A-50 | Sequence 50, Appl |
| 16 | 146 | 40.9 | 562 | 2 US-08-883-795A-38 | Sequence 38, Appl |
| 17 | 146 | 40.9 | 562 | 6 5185259-3 | Patent No. 5185259 |
| 18 | 146 | 40.9 | 562 | 6 5200340-2 | Patent No. 5200340 |
| 19 | 146 | 40.9 | 562 | 6 5344773-2 | Patent No. 5344773 |
| 20 | 141 | 39.5 | 655 | 1 US-08-148-910-12 | Sequence 12, Appl |
| 21 | 141 | 39.5 | 655 | 1 US-08-448-937A-12 | Sequence 12, Appl |
| 22 | 139 | 38.5 | 356 | 1 US-08-427-640-8 | Sequence 8, Appl |
| 23 | 134 | 37.5 | 477 | 2 US-08-560-098A-51 | Sequence 51, Appl |
| 24 | 131 | 36.7 | 138 | 2 US-08-797-689-12 | Sequence 12, Appl |
| 25 | 131 | 36.7 | 365 | 1 US-08-093-741-83 | Sequence 83, Appl |
| 26 | 131 | 36.7 | 365 | 1 US-08-720-012-83 | Sequence 83, Appl |
| 27 | 131 | 36.7 | 393 | 2 US-08-560-098A-44 | Sequence 44, Appl |

| | | | | | |
|----|-----|------|-----|---------------------|--------------------|
| 28 | 131 | 36.7 | 393 | 4 US-08-967-024C-24 | Sequence 24, Appl |
| 29 | 131 | 36.7 | 393 | 4 US-08-967-024C-25 | Sequence 25, Appl |
| 30 | 131 | 36.7 | 411 | 1 US-08-087-163-1 | Sequence 1, Appl |
| 31 | 131 | 36.7 | 411 | 1 US-08-286-748B-18 | Sequence 18, Appl |
| 32 | 131 | 36.7 | 411 | 1 US-08-153-799-19 | Sequence 19, Appl |
| 33 | 131 | 36.7 | 411 | 2 US-08-560-098A-48 | Sequence 48, Appl |
| 34 | 131 | 36.7 | 411 | 4 US-09-181-816-1 | Sequence 1, Appl |
| 35 | 131 | 36.7 | 430 | 1 US-07-942-157A-3 | Sequence 3, Appl |
| 36 | 131 | 36.7 | 430 | 6 5219569-2 | Patent No. 5219569 |
| 37 | 131 | 36.7 | 432 | 6 5188829-1 | Patent No. 5188829 |
| 38 | 131 | 36.7 | 432 | 2 US-08-560-098A-47 | Sequence 47, Appl |
| 39 | 128 | 35.9 | 355 | 1 US-08-427-640-6 | Sequence 6, Appl |
| 40 | 127 | 35.6 | 326 | 4 US-09-411-977-3 | Sequence 3, Appl |
| 41 | 127 | 35.6 | 347 | 2 US-08-811-949-1 | Sequence 1, Appl |
| 42 | 127 | 35.6 | 354 | 2 US-08-811-949-61 | Sequence 61, Appl |
| 43 | 127 | 35.6 | 355 | 1 US-08-137-116-1 | Sequence 1, Appl |
| 44 | 127 | 35.6 | 355 | 1 US-08-217-618-1 | Sequence 1, Appl |
| 45 | 127 | 35.6 | 355 | 1 US-08-427-640-2 | Sequence 2, Appl |

ALIGNMENTS

```
RESULT 1
US-09-411-977-2
; Sequence 2, Application US/09411977
; Patent No. 6372473
; GENERAL INFORMATION:
; APPLICANT: Moore, Paul A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: PR378P1
; CURRENT APPLICATION NUMBER: US/09/411,977
; CURRENT FILING DATE: 1998-10-04
; EARLIER APPLICATION NUMBER: 09/084,491
; EARLIER FILING DATE: 1998-05-27
; EARLIER APPLICATION NUMBER: 60/048,000
; EARLIER FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-411-977-2

Query Match      100.0%; Score 357; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 9.6e-36;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CFWDNGHLYREDQTSFAPGLRCLNWLDAQSLASAPVSGAGNHSYCRNPDDEPRGWCYV 60
DB      25 CFWDNGHLYREDQTSFAPGLRCLNWLDAQSLASAPVSGAGNHSYCRNPDDEPRGWCYV 84

RESULT 2
US-08-811-949-63
; Sequence 63, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIMMO, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUIT
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, WAIR & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
```

CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 472 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-63

Query Match
Best Local Similarity 44.9%; Score 150; DB 2; Length 472;
Matches 31; Conservative 4; Mismatches 22; Indels 12; Gaps 3;

Qy 1 CFWDNHLYREDQTSAPAGRLCLNWLDAQGLASAPVS-----GAGNHSYCRNPDE 51
Db 37 CYEDQGISYRGWTASTAGAECTNW--NSSALAQKPYSGRRPDIRLGLGHNHYCRNPDR 94

Qy 52 DPGPWCYV 60
Db 95 DSK-PWCYV 102

RESULT 3
US-08-811-949-2
Sequence 2, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NITAO, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.

REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 83 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-811-949-2

Query Match
Best Local Similarity 40.9%; Score 146; DB 2; Length 83;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

Qy 1 CFWDNHLYREDQTSAPAGRLCLNWLDAQGLASAPVS-----GAGNHSYCRNPDE 51
Db 1 CYEDQGISYRGWTASTAGAECTNW--NSSALAQKPYSGRRPDIRLGLGHNHYCRNPDR 58

Qy 52 DPGPWCYV 60
Db 59 DSK-PWCYV 66

RESULT 4
US-08-811-949-49
Sequence 49, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NITAO, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-49

Query Match 40.9%; Score 146; DB 2; Length 437;

Best Local Similarity 43.5%; Pred. No. 7.1e-10; Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3.

```

1      RESULT 5
2      US-08-811-949-51
3      : Sequence 51, Application US/08811949
4      : Patent No. 5840533
5      :
6      : GENERAL INFORMATION:
7      :
8      : APPLICANT: NINA, MINEO
9      : APPLICANT: SAITO, YOSHITAKA
10     : APPLICANT: SASAKI, HITOSHI
11     : APPLICANT: HAYASHI, MASAKO
12     : APPLICANT: NOTANI, JOUJI
13     : APPLICANT: KOBAYASHI, MASAKAZU
14     : TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
15     : NUMBER OF SEQUENCES: 67
16     :
17     : CORRESPONDENCE ADDRESSES:
18     : ADDRESSES: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
19     : STREET: 1765 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
20     : CITY: ARLINGTON
21     : STATE: VA
22     : COUNTRY: USA
23     :
24     : ZIP: 22202
25     :
26     : COMPUTER READABLE FORM:
27     :
28     : MEDIUM TYPE: floppy disk
29     : COMPUTER: IBM PC compatible
30     : OPERATING SYSTEM: PC-DOS/MS-DOS
31     : SOFTWARE: Patent In Release #1.0, Version #1.30
32     :
33     : CURRENT APPLICATION DATA:
34     : APPLICATION NUMBER: US/08/811,949
35     : FILING DATE: 05-MAR-1997
36     :
37     : CLASSIFICATION: 435
38     :
39     : ATTORNEY/AGENT INFORMATION:
40     :
41     : NAME: OBLON, NORMAN F.
42     : REGISTRATION NUMBER: 24,618
43     : REFERENCE/DOCKET NUMBER: 18-966-0
44     : TELECOMMUNICATION INFORMATION:
45     : TELEPHONE: 703-413-2220
46     : TELEFAX: 703-413-2220
47     :
48     : INFORMATION FOR SEQ ID NO: 51:
49     :
50     : SEQUENCE CHARACTERISTICS:
51     :
52     : LENGTH: 437 amino acids
53     : TYPE: amino acid
54     : TOPOLOGY: linear
55     :
56     : MOLECULE TYPE: protein
57     :
58     : US-08-811-949-51

```

| | | | | |
|--------------------------|--------|--------------------|------------|-------------|
| Query Match | 40.9% | Score 146; | DB 2; | Length 437; |
| Best Local Similarity | 43.5%; | Pred. No. 7.1e-10; | | |
| Matches 30; Conservative | 5; | Mismatches 22; | Indels 12; | Gaps 3 |

RESULT 6
US-08-811-949-55
: Sequence 55, Application US/08811949

Patent No.5840533
GENERAL INFORMATION:
APPLICANT: NIMA, MINMO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, WATER & NEUSTADT,
ADDRESSER: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release
CURRENT APPLICATION NUMBER: US/08/811,949
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
IS-08-811,949-55

| | | | | |
|-----------------------|--------|--------------------|-------|-----------------------------------|
| Query Match | 40.9%; | Score 146; | DB 2; | Length 437; |
| Best Local Similarity | 43.5%; | Pred. No. 7.1e-10; | | |
| Matches | 30; | Conservative | 5; | Mismatches 22; Indels 12; Gaps 3; |

RESULT 7
 US-08-811-949-57
 : Sequence 57, Application US/08811949
 : Patent No 584053
 : GENERAL INFORMATION:
 : APPLICANT: NIWA, MINEO
 : APPLICANT: SAITO, YOSHIMASA
 : APPLICANT: SASAKI, HIROSHI
 : APPLICANT: HAYASHI, MASAO
 : APPLICANT: NOTANI, JOUJI
 : APPLICANT: KOBAYASHI, MASAKAZU
 : TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
 : NUMBER OF SEQUENCES: 67
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: OBLOM, SIVYAK, MCLELLAND, MAIER & NEUSTADT
 : ADDRESS: P.C.
 : STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
 : CITY: ARLINGTON

STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-57

Query Match 40.9%; Score 146; DB 2; Length 437;
Best Local Similarity 43.5%; Pred. No. 7.1e-10;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWNGHLYREDQSPAPGLRCINWLDAGSLAPVS-----GAGNHSYCRNPDE 51
DB 2 CYBDQGISYRGWTASGAECTNM--NSSALAQKPYSGRRPDALRLGIGNHNYCRNPDR 59

QY 52 DPRGWCYV 60
DB 60 DSK-PWCYV 67

RESULT 8
US-07-609-510B-16
Sequence 16, Application US/07609510B
Patent No. 5326700
GENERAL INFORMATION:
APPLICANT: Berg et al.
TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue Pl
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: IN
COUNTRY: U.S.A.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/609,510B
FILING DATE: 19901106
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-609-510B-16

Query Match 40.9%; Score 146; DB 1; Length 527;
Best Local Similarity 43.5%; Pred. No. 8.9e-10;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWNGHLYREDQSPAPGLRCINWLDAGSLAPVS-----GAGNHSYCRNPDE 51
DB 92 CYBDQGISYRGWTASGAECTNM--NSSALAQKPYSGRRPDALRLGIGNHNYCRNPDR 149

QY 52 DPRGWCYV 60
DB 150 DSK-PWCYV 157

RESULT 9
US-08-811-949-39
Sequence 39, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NINA, MINO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JODI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-39

Query Match 40.9%; Score 146; DB 2; Length 527;
Best Local Similarity 43.5%; Pred. No. 8.9e-10;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWNGHLYREDQSPAPGLRCINWLDAGSLAPVS-----GAGNHSYCRNPDE 51
DB 92 CYBDQGISYRGWTASGAECTNM--NSSALAQKPYSGRRPDALRLGIGNHNYCRNPDR 149

QY 52 DPRGWCYV 60
DB 150 DSK-PWCYV 157

RESULT 10
PCT-US91-01025A-2


```
Sequence 2, Application PC/TUS9101025A
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSER: Genentech, Inc.
STREET: 460 Point San Francisco Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/01025A
FILING DATE: 19910214
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/486,657
FILING DATE: 1 March 1990
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 454P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/265-1896
TELEFAX: 415/952-9881
TRFEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
PCT-US91-01025A-2

Query Match          40.9%; Score 146; DB 5; Length 527;
Best Local Similarity 43.5%; Pred. No. 8.9e-10;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

Qy 1 CFMNGHLVREDQTSAPAGRLCLNWLDAQSGLASAPVS-----GAGNSYCRNPDE 51
   |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 92 CYEDQGISYRGTWSTAESGAECTNW--NSSALAKPYSGRRPDAIRLGLGNHYCRNPDR 149
   |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

Qy 52 DPGPWCYV 60
   |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 150 DSK-PWCYV 157

RESULT 11
5185259-8
Patent No. 5185259
APPLICANT: GOEDDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
VEHAR, GORDON A.
TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
ACTIVATOR
NUMBER OF SEQUENCES: 15
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/489,855
FILING DATE: 02-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 12,694
FILING DATE: 09-FEB-1987
APPLICATION NUMBER: 483,052
FILING DATE: 07-APR-1983
APPLICATION NUMBER: 398,003
FILING DATE: 14-JUL-1982
APPLICATION NUMBER: 374,860
FILING DATE: 05-MAY-1982
```

```
SEQ ID NO:8:
LENGTH: 527
5185259-8
Query Match          40.9%; Score 146; DB 6; Length 527;
Best Local Similarity 43.5%; Pred. No. 8.9e-10;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

Qy 1 CFMNGHLVREDQTSAPAGRLCLNWLDAQSGLASAPVS-----GAGNSYCRNPDE 51
   |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 92 CYEDQGISYRGTWSTAESGAECTNW--NSSALAKPYSGRRPDAIRLGLGNHYCRNPDR 149
   |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

Qy 52 DPGPWCYV 60
   |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 150 DSK-PWCYV 157

RESULT 12
5520913-1
Patent No. 5520913
APPLICANT: ANDERSON, STEPHEN; BENNETT, WILLIAM F.; BOTSTEIN,
DAVID; HIGGINS, DEBORAH L.; PAONI, NICHOLAS F.; ZOLLER, MARK J.
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR HAVING
ZYMOGENIC PROPERTIES
NUMBER OF SEQUENCES: 35
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/88,451
FILING DATE: 06-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 770,510
FILING DATE: 03-OCT-1991
APPLICATION NUMBER: 384,608
FILING DATE: 24-JUL-1989
APPLICATION NUMBER: 240,856
FILING DATE: 02-SEP-1988
SEQ ID NO:1:
LENGTH: 527
5520913-1

Query Match          40.9%; Score 146; DB 6; Length 527;
Best Local Similarity 43.5%; Pred. No. 8.9e-10;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

Qy 1 CFMNGHLVREDQTSAPAGRLCLNWLDAQSGLASAPVS-----GAGNSYCRNPDE 51
   |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 92 CYEDQGISYRGTWSTAESGAECTNW--NSSALAKPYSGRRPDAIRLGLGNHYCRNPDR 149
   |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

Qy 52 DPGPWCYV 60
   |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 150 DSK-PWCYV 157

RESULT 13
5200340-6
Patent No. 5200340
APPLICANT: FOSTER, DONALD C.; MULVHILL, EILEEN R.; O'HARA,
PATRICK J.; PINDEL, KURT; YOSHITAKE, SHINJI
TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
ACTIVATORS
NUMBER OF SEQUENCES: 34
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/53,412
FILING DATE: 22-MAY-1987
SEQ ID NO:6:
LENGTH: 546
5200340-6

Query Match          40.9%; Score 146; DB 6; Length 546;
Best Local Similarity 43.5%; Pred. No. 9.2e-10;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

Qy 1 CFMNGHLVREDQTSAPAGRLCLNWLDAQSGLASAPVS-----GAGNSYCRNPDE 51
   |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
```

Mon Apr 7 10:22:27 2003

us-10-057-951-2_copy_25_84.rai

Page 6

Db 127 CYEDQGISYRGWTSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLIGNNHYCRNPDR 184
Qy 52 DPGPWCYV 60
Db 185 DSK-PWCYV 192

RESULT 14
US-08-811-949-43
; Sequence 43, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIMA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 562 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-811-949-43

Query Match 40.9%; Score 146; DB 2; Length 562;
Best Local Similarity 43.5%; Pred. No. 9.6e-10;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;
Qy 1 CFWNDGHLTYREDQTSPPAGRLCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 51
Db 127 CYEDQGISYRGWTSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLIGNNHYCRNPDR 184
Qy 52 DPGPWCYV 60
Db 185 DSK-PWCYV 192

RESULT 15
US-08-560-098A-50
; Sequence 50, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENZEL, Stephan
; APPLICANT: HEINZEL-WITLAND, Regina
; APPLICANT: STEFFENS, Gerd Josef

; TITLE OF INVENTION: Proteins having fibrinolytic and
; COAGULATION-INHIBITING PROPERTIES
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Eversen, McKee, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 146/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8844
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 562 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-50

Query Match 40.9%; Score 146; DB 2; Length 562;
Best Local Similarity 43.5%; Pred. No. 9.6e-10;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;
Qy 1 CFWNDGHLTYREDQTSPPAGRLCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 51
Db 127 CYEDQGISYRGWTSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLIGNNHYCRNPDR 184
Qy 52 DPGPWCYV 60
Db 185 DSK-PWCYV 192

Search completed: April 7, 2003, 09:22:30
CPU time: 6.98753 secs

C:\Genetics:

A:Gene: CESP:P45E12.2
A:Introns: 55/1; 152/3; 392/2; 650/3; 691/3; 731/3
C:Superfamily: transcription initiation factor IIB 90K chain, transcription initiation

Query Match 9.6%; Score 87.5; DB 2; Length 759;
Best Local Similarity 30.4%; Pred. No. 4.2;
Matches 28; Conservative 17; Mismatches 32; Indels 15; Gaps 5;

QY 1 SGEAGVPEKPCEDL--RCPEPTTSQLPAFTTEIQEASEGPGADY-----OVFAPNAL 53
DB 577 ASESTIQKRSIFDLTEECSEFTSKNSPKYNLKVESAS--PSTSHVSSTIEHKFVP----- 630
QY 54 PARSEAAVQPIGTSQRVNMSEKDLGTL 85
DB 631 PARSRVAKVPIIGAKKALN--EVKNVHV 660

RESULT 3
A6510
leucyl tRNA synthetase [imported] - Chlamydia pneumoniae (strain J138)
C:Species: Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: A6510
R.Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A6510; PMID:2030349; PMID:10871362
A:Accession: A6510
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-820 <STO>
A:Cross-references: GB:BA000008; NID:g8978526; PIDN:BA98363.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: leuS
C:Superfamily: leucine-tRNA ligase

Query Match 9.4%; Score 86.5; DB 2; Length 820;
Best Local Similarity 24.1%; Pred. No. 5.7;
Matches 38; Conservative 28; Mismatches 49; Indels 43; Gaps 9;

QY 20 TTSQLPAFTTEIQE-----ASEGPGADVQVFAPNALPARSEAA--VQPIGTS 69
DB 235 TQSGLEAFTTRLDTLGVSFLVIAPEHPDLSIV-----SEQRDEVTAAYQESLRKS 268
QY 70 QRVNMS-KEKDLGLGY----VLGITMV-----IIAIGAGIILGYSYKRGDLKEQ 119
DB 289 ERDRISSVKRTGVTGNVAKHPITGNLPLVWISDVVLGVTGVVGV-----PA 339
QY 120 HDQKVCEREMQRTLPISAFNPTCEIVDEKTVVHTS 157
DB 340 HDER--DREFAEM-----FSLPIHEVIDDNGVCISHN 369

RESULT 4
C72113
leucine-tRNA ligase (EC 6.1.1.4) [similarity] - Chlamydia pneumoniae (strains CWL029
C:Species: Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 03-Jun-2002
C:Accession: C72113; F81557
R.Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; PMID:99206606; PMID:10192388
A:Accession: C72113
A:Molecule type: DNA
A:Residues: 1-820 <ARN>
A:Cross-references: GB:AE001602; GB:AE001363; NID:g4376416; PIDN:AA18306.1; PID:g437642
A:Experimental source: strain CWL029
R:Read, T.D.; Brinham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gilm, M.; Nelson, W.; Deboy, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.

A:Reference number: A81500; PMID:20150255; PMID:10684935
A:Accession: F81557
A:Molecule type: DNA
A:Residues: 1-820 <REA>
A:Cross-references: GB:AE002219; GB:AE002161; NID:g7189524; PIDN:AAJ8433.1; PID:g718952
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: leuS; CP0618
C:Superfamily: leucine-tRNA ligase
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 9.4%; Score 86.5; DB 2; Length 820;
Best Local Similarity 24.1%; Pred. No. 5.7;
Matches 38; Conservative 28; Mismatches 49; Indels 43; Gaps 9;

QY 20 TTSQLPAFTTEIQE-----ASEGPGADVQVFAPNALPARSEAA--VQPIGTS 69
DB 235 TQSGLEAFTTRLDTLGVSFLVIAPEHPDLSIV-----SEQRDEVTAAYQESLRKS 268
QY 70 QRVNMS-KEKDLGLGY----VLGITMV-----IIAIGAGIILGYSYKRGDLKEQ 119
DB 289 ERDRISSVKRTGVTGNVAKHPITGNLPLVWISDVVLGVTGVVGV-----PA 339
QY 120 HDQKVCEREMQRTLPISAFNPTCEIVDEKTVVHTS 157
DB 340 HDER--DREFAEM-----FSLPIHEVIDDNGVCISHN 369

RESULT 5
T31432
K-Cl cotransport protein 2, furosemide-sensitive - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 07-Dec-1999
C:Accession: T31432
R.Payne, U.A.; Stevenson, T.J.; Donaldson, L.F.
J. Biol. Chem. 271, 16245-16252, 1996
A:Title: Molecular characterization of a putative K-Cl cotransporter in rat brain: a ne
A:Reference number: Z21031; PMID:96279171; PMID:8663311
A:Accession: T31432
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1116 <PAY>
A:Cross-references: EMBL:U55816; NID:g1403708; PID:g1403709; PIDN:AAC52635.1
A:Experimental source: strain Sprague Dawley; clone ERB10; 5ERB12; brain
C:Genetics:
A:Gene: KCC2
C:Keywords: transmembrane protein

Query Match 9.3%; Score 85.5; DB 2; Length 1116;
Best Local Similarity 27.1%; Pred. No. 10;
Matches 32; Conservative 15; Mismatches 44; Indels 27; Gaps 4;

QY 79 KKDGLTGLVGLITMVIIAIGAGIILGYSY-----KRGDLKEQH--DQKVCEREMQ 130
DB 872 KKDGLTFLHLRTAEVEVEMHESDISAYTYEKLTVMEQSQILKQWHLTKEREKEIQ 931
QY 131 RIT-----LPLSAFTNPTC---EIVDEKTVVHTSQTVPDPOGSGTP 169
DB 932 SITDSRGSIRKRPANTRRLNVPETACDNEKPEEYQGLHDSAPSCSSPSP 989

RESULT 6
B82525
primosomal protein N' XP2689 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: B82525
R.randomous. The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; PMID:20365717; PMID:10910347
A:Note: For a complete list of authors see reference number A59328 below
A:Accession: B82525

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-733 <SIM>
A;Cross-references: GB:AE004075; GB:AE003849; NID:99107929; PIDN:AAF65486.1; GSPDB:GN001
A;Experimental source: strain 9asc
A;Authors: A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Brienes, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carriao, D.M.; Carter, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.U.S.
Submitted to Genbank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kurame, E.E.; Laigz
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, B.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshahro, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XRF2689

Query Match 9.1%; Score 83.5; DB 2; Length 733;
Best Local Similarity 22.9%; Pred. No. 9.6; Indels 49; Gaps 3;
Matches 35; Conservative 17; Mismatches 52;

Db 9 KRPCEDLRCPETTSQALPAFTTEIOEASE-----GPGADEVOVAPAN 51
181 KKPTTVALIQTTSNTPVANTEQKAVETLNNVGFKTYLLDGVSGTEVYLQALIN 240

Qy 52 ALPARSAAAVPVYIGISQVRNMSKEKKDL-----GTLGYV 88
241 CLAAKQALVLPETIGTLTNFARLIGIPVHALSKLADNERARVMAARGEARLV 300

Db 89 LGITMYIILAIAGIILGYSYKRGKDLKEQHD 121
301 LGTRSAVFTLPHAGLIT-----VDEEHD 324

RESULT 7
T25592
hypothetical protein C32E12.4 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T25592
R;Wilson, L.
Submitted to the EMBL Data Library, November 1996
A;Description: The sequence of C. elegans cosmid C32E12.
A;Reference number: Z20055
A;Accession: T25592
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1624 <MIL>
A;Cross-references: EMBL:U80032; PIDN:AA55880.1; GSPDB:GN00019; CESP:C32E12.4
A;Experimental source: strain Bristol N2; clone C32E12
C;Genetics:
A;Gene: CESP:C32E12.4
A;Map position: 1
A;Intons: 12/2; 123/2; 175/1; 249/3; 325/3; 357/3; 513/3; 570/1; 617/3; 674/3; 694/1; 8

Query Match 9.1%; Score 83; DB 2; Length 1624;
Best Local Similarity 18.1%; Pred. No. 27;
Matches 36; Conservative 33; Mismatches 76; Indels 54; Gaps 5;

Qy 3 EAGVPERKPCEDLRCP-----TTSQALPAFTTEIOEASEGGADEVQV 46
Db 394 KSSVEETKTETESPPVQATPEETEKIDFGSTPVTAVERGVNVLVTSSEDEEDVEI 453
Qy 47 FAPANALPARSAAAVPVYIGISQVRNMSKEKKDLGTLGYVIGITMYIILAIAGIIL 106
Db 454 IKRIBQ-PEKASVLPVPIKVVSSIKINGIEYKDV----- 488
Qy 107 GYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPTCEIYDEKTVV---HTSQTPV- 161

Db 469 -----EKTKEDKKEVDKSVKVAATNIDVPTPIPTSETSKVPEMLISSAVKVLNDHGSIPEH 544
Qy 162 ----DPEGSTPLMQOACT 176
Db 545 EKSDDPQEFSTIEKSESST 563

RESULT 8
T45031
hypothetical protein Y396B.e [imported] - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C;Accession: T45031
R;Wilson, R.; Ainscough, R.; Anderson, K.; Baynes, C.; Berke, M.; Bonfield, J.; Burton,
raser, A.; Fulton, L.; Gardner, A.; Green, P.; Hawtins, T.; Hillier, L.; Jier, M.; John
B.; O'Callaghan, M.; Parsons, J.; Percy, C.; Rifken, L.; Roopra, A.; Saunders, D.
Nature 368, 32-38, 1994
A;Authors: Showkneen, R.; Sims, M.; Smaildon, N.; Smith, A.; Smith, M.; Sonhammer, E.; S
lock, L.; Wilkinson-Sproat, J.; Wohlman, P.
A;Title: 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.
A;Reference number: S43531; MUID:94150718; PMID:7906398
A;Accession: T45031
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1365 <MIL>
A;Cross-references: EMBL:AL12896; NID:96434440; PIDN:CAB60910.1; PID:96434445
A;Experimental source: clone Y396B
C;Genetics:
A;Map position: 3
A;Intons: 10/1; 37/1; 171/2; 844/2; 1074/2; 1115/3; 1194/2; 1233/3
A;Note: Y396B.e

Query Match 8.9%; Score 81.5; DB 2; Length 1365;
Best Local Similarity 23.7%; Pred. No. 30; Indels 11; Gaps 3;
Matches 28; Conservative 22; Mismatches 57;

Qy 18 PETTSQALPAFTTEIOEASEGPGADEVOVAPANALPARSAAAVPVYIGISQVRNMSK 77
Db 785 PQNSGCVNTGTGNGNGSGPSGDTGT--GPTSGTGTAAGV-----SSVSTVSD 837

Qy 78 EKDKLGTGYVIGITMYIILAIAGIILGY---SYKRGKDLKEQHDQKVCEREMOR 131
Db 838 ESSDPILFVIVPVLLIVLIVIGIVFIWYKREKREKREKAEENEKLIAREERMR 895

RESULT 9
H64568
histidine kinase - *Helicobacter pylori* (strain 26695)
C;Species: *Helicobacter pylori*
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 20-Apr-2000
C;Accession: H64568
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Kalk, H.G.; Glodek, A.; McKen
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathey, L.
Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser, C.
A;Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: H64568
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-803 <TCM>
A;Cross-references: GB:AE000555; GB:AE000511; NID:g2313485; PIDN:AA07457.1; PID:g231345
C;Superfamily: Helicobacter pylori histidine kinase; response regulator homology
F/679-792/Domain: response regulator (Asp) (covalent) #status predicted
F/729/Binding site: phosphate (Asp)

Query Match 8.8%; Score 80.5; DB 2; Length 803;
Best Local Similarity 25.2%; Pred. No. 20; Indels 27; Gaps 5;
Matches 36; Conservative 22; Mismatches 58;

| | | | | |
|---|---|-------------------|---|-----|
| Oy | 31 | EIOEASBPADDEVQFPA | NALPARSEAAAVQPIGISQFVMSKKEKDLGTL--GYV | 88 |
| | | | | |
| Db | 227 | EVTPTKEFPPKPKHETX | KAKADTEENKA---PSIGVQYRVADVRLDHLMLNIGELV | 282 |
| | | | | |
| Oy | 89 | LGITMMVITIIAGAGII | IGYSYKRGKDLKEQHDQVCEREMQRTLPISATNPCEIYD | 148 |
| | | | | |
| Db | 283 | LGNKRLIRI----- | YS---DVERDYDGEKFLPEELNQVSSISAVT-----TD | 321 |
| | | | | |
| Oy | 149 | EKTVVVHTSQTPVPD | PEQEGSTPLM | 171 |
| | | | | |
| Db | 322 | LQLAVMKTRMQDVGK | VFNKKPRM | 344 |
| | | | | |
| RESULT 10 | | | | |
| G97438 | | | | |
| glutamate-cysteine ligase precursor, chloroplast (gamma-glutamylcysteine synthetase) (G97438) | | | | |
| C/Species: | Agrobacterium tumefaciens | | | |
| C/Date: | 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002 | | | |
| C/Accession: | G97438 | | | |
| R/Goodman, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quirillo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science | 294, 2223-2328, 2001 | | | |
| A/Title: | Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens | | | |
| A/Reference number: | A97359; PMID:11743194 | | | |
| A/Accession: | G97438 | | | |
| A/Status: | Preliminary | | | |
| A/Molecule type: | DNA | | | |
| A/Residues: | 1-457 <KUR> | | | |
| A/Cross-references: | GB:AB007869; PIDN:AKK86464.1; PID:G15155610; GSPDB:GN00169 | | | |
| C/Genetics: | | | | |
| A/Gene: | AGR_C_1167 | | | |
| Map position: circular chromosome | | | | |

```

Query Match          8.6%; Score 78.5; DB 2; Length 457;
Best Local Similarity 32.7%; Pred. No. 16;
Matches      33; Conservative    13; Mismatches   26; Indels    29; Gaps     6;

OY      2 GEAGVPEKRPCEDLRCPETTSQALPFTTEI---QEASEGPGA-----DEVGFAPAN 51
       ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      325 GADGGPMWRIC-----ALPAFWGLLYNGEADDAADATPADSPDE--VTALRN 371
       ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY      52 ALPARSEAAAV--QPVIIGISQRY----RNSKEKDGLTGG 86
       :|::| |:| ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      372 AVPAKGLAAEIKGKPLLGARQVDLSRTGLNRRRLNGEG 412
       :|::| |:| ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 11
AB2657
glutamate-cysteine ligase gshl [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C.Species: Agrobacterium tumefaciens
C.Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C.Accession: AB2657
R.Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.;
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutlyavin, T.; Levy, R.; Li, M.; McCelligott,
J.; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A.Authors: Xoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krepsan, W.; Perry, M.; Gordon-Kamm,
ster, E.M.
A.Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A.Reference number: AB2577; PMID:11743193
A.Accession: AB2657
A.Status: Preliminary
A.Molecule type: DNA
A.Residues: 1-457 <KDR>
A.Cross-references: GB:AEO08688; PIDN:ALA1672.1; PID:G17739016; GSFD:B:GN00186
A.Experimental source: strain C58 (Dupont)
C.Genetics:
A.Gene: gshl
A.Map position: circular chromosome

Query Match          8.6%; Score 78.5; DB 2; Length 457;
Best Local Similarity 32.7%; Pred. No. 16;
Matches      33; Conservative    13; Mismatches   26; Indels    29; Gaps     6;
```

QY 2 GEAIVPEKRPEDRCETTSQALPAFTTEI-----QASESGGA-----DEVVFPAN 51
 Db 325 GADGSPMRIC-----ALPAFWVGLLYNQEALDAADALTADWSFDE--VIALRN 371
 QY 52 ALPARSEAAV--OPVIGISRV-----RMSKKEKDGITGG 86
 Db 372 AVPAKGLMAETAGKPLGLIGIAQVLDISTRTGLAKRKRLANGEG 412

```

RESULT 12
A32555
major merizoite surface antigen precursor - Plasmodium chabaudi adami (fragment)
C/Species: Plasmodium chabaudi adami
C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 11-Jan-2000
C/Accession: A32555
R/Lew, A.M.; Langford, C.J.; Anders, R.F.; Kemp, D.J.; Saul, A.; Fardoulis, C.; Geysen,
Proc. Natl. Acad. Sci. U.S.A. 86, 3768-3772, 1989
A>Title: A protective monoclonal antibody recognizes a linear epitope in the precursor
A/Reference number: A32555; MUID:89264504; PMID:2471191
A/Accession: A32555
A/Molecule type: mRNA
A/Residues: 1-478 <LEM>
A/Cross-references: GB:J04568; NID:g160419; PID:g552206
C/Superfamily: major merizoite surface antigen
C/Keywords: glycoprotein; surface antigen
F119, 299/Binding site: carbohydrate (Asn) (covalent) #stratus predicted

```

| | | | | |
|-----------------------|------------------|----------------|------------|-------------|
| Query Match | 8.6%; | Score 78.5; | DB 2; | Length 478; |
| Best Local Similarity | 28.7%; | Pred. No. 17; | | |
| Matches 29; | Conservative 11; | Mismatches 38; | Indels 23; | Gaps 3; |

```

QY      3  EADVPRKCRDCELRCPETTSOAL-----AFTTIOQASEBPGADV-----44
      |||||
Db      266  EAQPPSAAQGE-ATTETTTTAAEPTPTKCGASTNKSSEFSTRAPAASTPEVAPSPAT 324
      |||||
QY      45  -----QVFAPANLTPARSAALVQVITISQRYMNSKKED 81
      |||||
Db      325  PAAPSSASPPAQPPAQPPVTSQPSQSESTNNVGSTQQVKA 365
      |||||

```

```

RESULT 13
T22759
hypothetical protein F55H12.3 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Aug-2000
C/Accession: T22759
R/Dobson, R.
submitted to the EMBL Data Library, October 1996
A/Reference number: Z19610
A/Accession: T22759
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-2824 <WIL>
A/Cross-references: EMBL;Z81091; PIDB:CA03143.1; GSPDB:GN00019; CESP:F55H12.3
A/Experimental source: clone F55H12
C/Genetics:
A/Gene: CESP:F55H12.3
A/Map position: 1
A/Intons: 87/1; 98/1; 126/2; 201/3; 343/3; 406/1; 576/3; 656/1; 825/3; 869/1; 909/1; 966/1; 1000/1; 1001/1; 1002/1; 1003/1; 1004/1; 1005/1; 1006/1; 1007/1; 1008/1; 1009/1; 1010/1; 1011/1; 1012/1; 1013/1; 1014/1; 1015/1; 1016/1; 1017/1; 1018/1; 1019/1; 1020/1; 1021/1; 1022/1; 1023/1; 1024/1; 1025/1; 1026/1; 1027/1; 1028/1; 1029/1; 1030/1; 1031/1; 1032/1; 1033/1; 1034/1; 1035/1; 1036/1; 1037/1; 1038/1; 1039/1; 1040/1; 1041/1; 1042/1; 1043/1; 1044/1; 1045/1; 1046/1; 1047/1; 1048/1; 1049/1; 1050/1; 1051/1; 1052/1; 1053/1; 1054/1; 1055/1; 1056/1; 1057/1; 1058/1; 1059/1; 1060/1; 1061/1; 1062/1; 1063/1; 1064/1; 1065/1; 1066/1; 1067/1; 1068/1; 1069/1; 1070/1; 1071/1; 1072/1; 1073/1; 1074/1; 1075/1; 1076/1; 1077/1; 1078/1; 1079/1; 1080/1; 1081/1; 1082/1; 1083/1; 1084/1; 1085/1; 1086/1; 1087/1; 1088/1; 1089/1; 1090/1; 1091/1; 1092/1; 1093/1; 1094/1; 1095/1; 1096/1; 1097/1; 1098/1; 1099/1; 1100/1; 1101/1; 1102/1; 1103/1; 1104/1; 1105/1; 1106/1; 1107/1; 1108/1; 1109/1; 1110/1; 1111/1; 1112/1; 1113/1; 1114/1; 1115/1; 1116/1; 1117/1; 1118/1; 1119/1; 1120/1; 1121/1; 1122/1; 1123/1; 1124/1; 1125/1; 1126/1; 1127/1; 1128/1; 1129/1; 1130/1; 1131/1; 1132/1; 1133/1; 1134/1; 1135/1; 1136/1; 1137/1; 1138/1; 1139/1; 1140/1; 1141/1; 1142/1; 1143/1; 1144/1; 1145/1; 1146/1; 1147/1; 1148/1; 1149/1; 1150/1; 1151/1; 1152/1; 1153/1; 1154/1; 1155/1; 1156/1; 1157/1; 1158/1; 1159/1; 1160/1; 1161/1; 1162/1; 1163/1; 1164/1; 1165/1; 1166/1; 1167/1; 1168/1; 1169/1; 1170/1; 1171/1; 1172/1; 1173/1; 1174/1; 1175/1; 1176/1; 1177/1; 1178/1; 1179/1; 1180/1; 1181/1; 1182/1; 1183/1; 1184/1; 1185/1; 1186/1; 1187/1; 1188/1; 1189/1; 1190/1; 1191/1; 1192/1; 1193/1; 1194/1; 1195/1; 1196/1; 1197/1; 1198/1; 1199/1; 1200/1; 1201/1; 1202/1; 1203/1; 1204/1; 1205/1; 1206/1; 1207/1; 1208/1; 1209/1; 1210/1; 1211/1; 1212/1; 1213/1; 1214/1; 1215/1; 1216/1; 1217/1; 1218/1; 1219/1; 1220/1; 1221/1; 1222/1; 1223/1; 1224/1; 1225/1; 1226/1; 1227/1; 1228/1; 1229/1; 1230/1; 1231/1; 1232/1; 1233/1; 1234/1; 1235/1; 1236/1; 1237/1; 1238/1; 1239/1; 1240/1; 1241/1; 1242/1; 1243/1; 1244/1; 1245/1; 1246/1; 1247/1; 1248/1; 1249/1; 1250/1; 1251/1; 1252/1; 1253/1; 1254/1; 1255/1; 1256/1; 1257/1; 1258/1; 1259/1; 1260/1; 1261/1; 1262/1; 1263/1; 1264/1; 1265/1; 1266/1; 1267/1; 1268/1; 1269/1; 1270/1; 1271/1; 1272/1; 1273/1; 1274/1; 1275/1; 1276/1; 1277/1; 1278/1; 1279/1; 1280/1; 1281/1; 1282/1; 1283/1; 1284/1; 1285/1; 1286/1; 1287/1; 1288/1; 1289/1; 1290/1; 1291/1; 1292/1; 1293/1; 1294/1; 1295/1; 1296/1; 1297/1; 1298/1; 1299/1; 1300/1; 1301/1; 1302/1; 1303/1; 1304/1; 1305/1; 1306/1; 1307/1; 1308/1; 1309/1; 1310/1; 1311/1; 1312/1; 1313/1; 1314/1; 1315/1; 1316/1; 1317/1; 1318/1; 1319/1; 1320/1; 1321/1; 1322/1; 1323/1; 1324/1; 1325/1; 1326/1; 1327/1; 1328/1; 1329/1; 1330/1; 1331/1; 1332/1; 1333/1; 1334/1; 1335/1; 1336/1; 1337/1; 1338/1; 1339/1; 1340/1; 1341/1; 1342/1; 1343/1; 1344/1; 1345/1; 1346/1; 1347/1; 1348/1; 1349/1; 1350/1; 1351/1; 1352/1; 1353/1; 1354/1; 1355/1; 1356/1; 1357/1; 1358/1; 1359/1; 1360/1; 1361/1; 1362/1; 1363/1; 1364/1; 1365/1; 1366/1; 1367/1; 1368/1; 1369/1; 1370/1; 1371/1; 1372/1; 1373/1; 1374/1; 1375/1; 1376/1; 1377/1; 1378/1; 1379/1; 1380/1; 1381/1; 1382/1; 1383/1; 1384/1; 1385/1; 1386/1; 1387/1; 1388/1; 1389/1; 1390/1; 1391/1; 1392/1; 1393/1; 1394/1; 1395/1; 1396/1; 1397/1; 1398/1; 1399/1; 1400/1; 1401/1; 1402/1; 1403/1; 1404/1; 1405/1; 1406/1; 1407/1; 1408/1; 1409/1; 1410/1; 1411/1; 1412/1; 1413/1; 1414/1; 1415/1; 1416/1; 1417/1; 1418/1; 1419/1; 1420/1; 1421/1; 1422/1; 1423/1; 1424/1; 1425/1; 1426/1; 1427/1; 1428/1; 1429/1; 1430/1; 1431/1; 1432/1; 1433/1; 1434/1; 1435/1; 1436/1; 1437/1; 1438/1; 1439/1; 1440/1; 1441/1; 1442/1; 1443/1; 1444/1; 1445/1; 1446/1; 1447/1; 1448/1; 1449/1; 1450/1; 1451/1; 1452/1; 1453/1; 1454/1; 1455/1; 1456/1; 1457/1; 1458/1; 1459/1; 1460/1; 1461/1; 1462/1; 1463/1; 1464/1; 1465/
```

```
QY      11 PCDELT-----RCPEPTTSOALPFAETTEIGESBEGCAGDEVGFAPANALLPARSEAAA 61
          ||| |||||
Db     2590 PCSULSENATISIVTQCSTCSDAIPFAGNCNCLQQNGKSS-----MIRNCLPF-----E 2638
          ||| |||||
QY      62 VQPVITGISQRVRANKSEKKDLGTGLGVVLGITMMVITIIAIGAILIGSYIKSGKDILKE-QH 120
```

Db 2639 VTFVPEPNSNAIKI-----VLGVFVGLIILIIIVLVCRKQIIMFRKTDPSDQH 2689
 Qy 121 DQKVCEREMORITLPLSAFTNPTCEIYDEKTVVHHNSQFVPOQESRP 169
 Db 2690 -----VALSHMDNATNRNEBENONPTSTVYTRIPPOAPLE 2725

RESULT 14

AD0760
 diol dehydratase medium chain [imported] - Salmonella enterica subsp. enterica serovar Typhimurium
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
 C:Accession: AD0760
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium
 A:Reference number: AB0502; PMID:11677608
 A:Accession: AD0760
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-224 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD02402.1; PID:G16503272; GSPDB:GN00176
 C:Genetics:
 A:Gene: pdu

Query Match 8.5%; Score 78; DB 2; Length 224;
 Best Local Similarity 27.0%; Pred. No. 7.9;
 Matches 34; Conservative 15; Mismatches 37; Indels 40; Gaps 6;

Qy 18 PFTTSQALPAFTTTEQASBEGADDEVQVFPANALPARSEAAVOPVIGISQVRMSK 77
 Db 37 PQTAAPAGSGFLTEVGEARQGTQDEVTI-----AVGPARGLAQTIVNIGL 82

Qy 78 EKXDLGTLGYL-----GITMVI-----IAT-----GAGIILGYSYRGRDL 116
 Db 83 PKKSI--LREVIAGIEBGRIVRVRPKSSDVAFAVEGNRLSGGISIGI--QSKDT 137

Qy 117 KEQHDQ 122
 Db 138 TVIHQ 143

RESULT 15

D98331
 oppd protein (AF305387) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)

C:Species: Agrobacterium tumefaciens
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
 C:Accession: D98331
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quicello, B.; Goldman, A.; Liu, F.; Wollem, C.; Allinger, M.; Dougherty, D.; Scott, C.; Dappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
 A:Reference number: A97359; PMID:11743194
 A:Accession: D98331
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-267 <KDR>
 A:Cross-references: GB:AE007870; PIDN:AAK90174.1; PID:G15160177; GSPDB:GN00170
 C:Genetics:
 A:Gene: AGR_1_3200
 A:Map position: linear chromosome

Query Match 8.5%; Score 77.5; DB 2; Length 267;
 Best Local Similarity 21.1%; Pred. No. 11;
 Matches 30; Conservative 24; Mismatches 43; Indels 45; Gaps 4;

Qy 48 APANALPARSEAA-----AVOPVIGISQVRMSKSKDLGTLGVLGITMNV 95
 Db 138 TVIHQ 143

Db 53 AINGLPARAQVSGDISWRDGLFTGRPMWG-----RDIGTIFQDTGATLNP 100
 Qy 96 II-----IAGGIIILGYSYRGRDLKEQ-----HDQKVCEREMORITL 134
 Db 101 VLTIGQVABEGAVRHLGLSMRQGRDLRLERVLPHPSHLISVPHQLSGQRQVVAI 160
 Qy 135 PLSAFTNPTCEIYDEKTVVHT 156
 Db 161 AAALARPALITADEATSAIDT 182

Search completed: April 7, 2003, 09:19:23
 Job time : 17.513 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:24 ; Search time 8.18711 Seconds
(without alignments)
906.823 Million cell updates/sec

Title: US-10-057-951-2_COPY_85_263

Sequence: 1 SEBAGVPEKRPCEDLRCPET.....PVDQEGSTPLMGQAQTPGA 179

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | ID | Description |
|------------|-------|--------------------|--------|-------------|
| 1 | 86.5 | 9.4 | 820 1 | SYL_CHLPP |
| 2 | 81 | 8.8 | 499 1 | GSHR_PLAT7 |
| 3 | 81 | 8.8 | 499 1 | GSHR_PLATK |
| 4 | 80 | 8.7 | 1238 1 | UAG2_HUMAN |
| 5 | 77.5 | 8.5 | 1709 1 | SN_HUMAN |
| 6 | 76.5 | 8.4 | 425 1 | TR16_RAT |
| 7 | 76.5 | 8.4 | 501 1 | GYG2_HUMAN |
| 8 | 76 | 8.3 | 505 1 | SPRO_SYNY3 |
| 9 | 75 | 8.2 | 595 1 | TNR8_HUMAN |
| 10 | 74.5 | 8.1 | 747 1 | YRGF_ECOLI |
| 11 | 73 | 8.0 | 260 1 | MYKB_METEX |
| 12 | 73 | 8.0 | 536 1 | YENI_SCHPO |
| 13 | 73 | 8.0 | 676 1 | ICP0_HSVBU |
| 14 | 73 | 8.0 | 698 1 | EF6_DEIRA |
| 15 | 72.5 | 7.9 | 400 1 | B3AR_MOUSE |
| 16 | 72.5 | 7.9 | 458 1 | DESM_XENLA |
| 17 | 72 | 7.9 | 646 1 | MU18_HUMAN |
| 18 | 72 | 7.9 | 952 1 | IF41_YEAST |
| 19 | 72 | 7.9 | 1781 1 | AKAC_HUMAN |
| 20 | 71.5 | 7.8 | 145 1 | YD54_METLA |
| 21 | 71.5 | 7.8 | 755 1 | RRE1_HUMAN |
| 22 | 71.5 | 7.8 | 817 1 | HUMB_MUSDO |
| 23 | 70.5 | 7.7 | 241 1 | MI41_SARMT |
| 24 | 70.5 | 7.7 | 272 1 | IBP1_RAT |
| 25 | 70.5 | 7.7 | 442 1 | BNE_DROME |
| 26 | 70.5 | 7.7 | 534 1 | FM2_ACTNA |
| 27 | 70.5 | 7.7 | 1001 1 | IR2_SYNY3 |
| 28 | 70.5 | 7.7 | 1182 1 | HAIR_MOUSE |
| 29 | 70.5 | 7.7 | 2278 1 | FAB1_YEAST |
| 30 | 70 | 7.6 | 310 1 | SUCA_DICDI |
| 31 | 70 | 7.6 | 393 1 | SDY1_RAT |
| 32 | 70 | 7.6 | 652 1 | CD93_HUMAN |
| 33 | 70 | 7.6 | 1505 1 | AT7B_SHEEP |

| | | | | | |
|----|------|-----|--------|------------|---------------------|
| 34 | 69.5 | 7.6 | 241 1 | MI41_SARMT | 008658 sarcocystis |
| 35 | 69.5 | 7.6 | 272 1 | IBP1_MOUSE | P47876 mus musculus |
| 36 | 69.5 | 7.6 | 358 1 | VCOM_ADR40 | P48753 human adeno |
| 37 | 69.5 | 7.6 | 388 1 | NHAA_ECOLI | P13778 escherichia |
| 38 | 69.5 | 7.6 | 513 1 | ATPA_MYCPD | Q28957 mycoplasma |
| 39 | 69.5 | 7.6 | 579 1 | PR17_HUMAN | O60508 homo sapien |
| 40 | 69.5 | 7.6 | 626 1 | PM17_MOUSE | O60536 mus musculu |
| 41 | 69.5 | 7.6 | 1025 1 | DPYD_PIG | Q28943 sus scrofa |
| 42 | 69 | 7.5 | 322 1 | NUCA_CAERL | P53596 caenorhabdi |
| 43 | 69 | 7.5 | 382 1 | ERG2_TREPA | Q9K474 bacillus ha |
| 44 | 69 | 7.5 | 683 1 | ERG2_TREPA | O83454 treponema p |
| 45 | 69 | 7.5 | 766 1 | METE_YEAST | P05694 saccharomyc |

ALIGNMENTS

RESULT 1
ID SYL_CHLPP STANDARD, PRT, 820 AA.
AC Q92930; Q9JQ86; 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leucyl-tRNA synthetase (EC 6.1.1.4) (leucine--tRNA ligase) (LeuRS).
GN LEU5 OR CPN0153 OR CP0618.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CW1029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.,
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA White T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA Linher K., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Gwin M., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mohn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=2030349; PubMed=10871362;
RA Shiba T., Ishii K., Hattori M., Tabuchi M., Kishi F., Ouchi K.,
RA Shida T., Ishii K., Hattori M., Kohara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CW1029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -1- CATALYTIC ACTIVITY: ATP + L-leucine + tRNA(Leu) = AMP +
CC diphosphate + L-leucyl-tRNA(Leu).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001602; AAD18306.1; --
CC EMBL: AB002219; AAF38433.1; --

DR EMBL: AP002545; BAA98363.1; -.
DR TIGR: CP0618; -.
DR InterPro: IPR002302; Leu-tryptophan.
DR InterPro: IPR002300; trna-tryptophan.
DR InterPro: IPR001412; trna-tryptophan.
DR Pfam: PF00133; trna-tryptophan.
DR PRINTS: PR00985; TRNASYNTHLEU.
DR TIGRfam: TIGR00396; Leu-bact.
DR PROSITE: PS00178; AA-TRNA-LIGASE-I; 1.
KM Anticodon-tryptophan synthetase; Protein biosynthesis; Ligase; ATP-binding;
FT SITE 40 51 "HIGH" REGION.
FT BINDING 601 605 "RMSKS" REGION.
FT ACT SITE 604 604 ATP (BY SIMILARITY).
SQ SEQUENCE 820 AA; 93965 MW; 520369FC098F1926 CRC64;

Query Match 9.4%; Score 86.5; DB 1; Length 820;
Best Local Similarity 24.1%; Pred. No. 2.9;
Matches 38; Conservative 28; Mismatches 49; Indels 43; Gaps 9;

QY 20 TTSGALPFTTEIQE-----ASEGPADEVQVAPANALPARSEAA-VQPVIGIS 69
Db 235 TQESLEAFTTDLTLGVSFLVIAPEHDLDSIV-----SEGRDEVTAVQESLRKS 286
QY 70 QRVAMNS-KKKDLGLGY---VLGITMV-----IIAIGAGIILGYSTRKDKLEQ 119
Db 289 ERDISSVKKTKGTGTGNVAKHPIITGNLLPWIISDVVLGVTGVWGV-----PA 339
QY 120 HDQKVCEREMQRTLPPLAFTNPTCEIVDEKTVVHTS 157
Db 340 HDER--DREFAE-----FSLPIHEVIDNGVCISHN 369

RESULT 2
GSHR_PLAF7 STANDARD; PRT; 499 AA.
ID GSHR_PLAF7
AC 015770;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glutathione reductase (EC 1.6.4.2) (GR) (GRASE).
GN GR1.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_Taxid=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Silberger T.-W., Walter R.D., Mueller S.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAINTAIN HIGH LEVELS OF REDUCED GLUTATHIONE IN THE
CYTOSOL. (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: NADPH + oxidized glutathione = NADP(+) + 2
glutathione.
CC -1- COFACTOR: FAD (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- MISCELLANEOUS: THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE
OXIDOREDUCTASES CLASS-I.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF027825; AAB84117.1; -.
DR HSPF: P00390; 1ALG.
DR InterPro: IPR001327; FAD pyr. redox.
DR InterPro: IPR000815; Hg_reductase.
DR InterPro: IPR001100; Pyr_redox.
DR InterPro: IPR000103; Pyridine_redox_2.

DR InterPro: IPR004099; Pyr_redox_dim.
DR Pfam: PF00070; pyr_redox_1.
DR Pfam: PF02852; pyr_redox_dim; 1.
DR PRINTS: PR00368; FADPNR.
DR PRINTS: PR00945; HGRDPTASE.
DR PRINTS: PR00411; PNDRTASEII.
DR PRINTS: PR00469; PNDRTASEII.
DR PRODOM: PD000139; FAD pyr. redox; 2.
DR PROSITE: PS00076; PYRIDINE REDOX 1; 1.
KM Redox-active center; Oxidoreductase; Flavoprotein; FAD; NADP.
FT INT_MET 0 0
FT NP_BIND 3 33 FAD (ADP PART) (POTENTIAL).
FT DISULFID 39 44 REDOX-ACTIVE (BY SIMILARITY).
FT NP_BIND 301 311 FAD (FLAVIN PART) (BY SIMILARITY).
FT ACT SITE 484 484 BY SIMILARITY.
SQ SEQUENCE 499 AA; 56288 MW; 43CCE0251E7B8244 CRC64;

Query Match 8.8%; Score 81; DB 1; Length 499;
Best Local Similarity 24.5%; Pred. No. 5.3;
Matches 34; Conservative 25; Mismatches 64; Indels 16; Gaps 7;

QY 29 TTEIQASEGPADEVQVAPANALPARSEAAVQPVIGISQVR---MNSKEKDLGT 84
Db 123 TKNNKNDNGPLNEE--ILEGRNIIAIVGNKVPFPVAGINTISSDFEYKESKIGI 180
QY 85 LGY-VLGITMMVITIIAIGAGIILGYSTRKDKLEHQ---KVCEREMQRTLPPLAFT 140
Db 181 VSGGYIAVELIIVIKRLG--IDSYIFARGNRIILKFDSEVINVLNDKNNINIVTFA 237
QY 141 NPTCEI--VDEKTVVHTS 157
Db 238 D-VVEIKKVSDBKLSIHL 255

RESULT 3
GSHR_PLAFK STANDARD; PRT; 499 AA.
ID GSHR_PLAFK
AC 094655;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glutathione reductase (EC 1.6.4.2) (GR) (GRASE).
GN GR2.
OS Plasmodium falciparum (isolate K1 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_Taxid=5839;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=96370813; PubMed=8774709;
FA Faerber P.M., Becker K., Mueller S.;
RL "Molecular cloning and characterization of a putative glutathione
reductase gene, the pGR2 gene, from Plasmodium falciparum.";
RL Eur. J. Biochem. 239:655-661(1996).
RN [2]
RP SEQUENCE OF 1-15, AND CHARACTERIZATION.
RX MEDLINE=96202957; PubMed=8613152;
FA Krutuh-Siegel R.L., Muller J.G., Lotzspeich F., Schirmer R.H.;
RL "Glutathione reductase and glutamate dehydrogenase of Plasmodium
falciparum, the causative agent of tropical malaria.";
RL Eur. J. Biochem. 235:345-350(1996).
CC -1- FUNCTION: MAINTAIN HIGH LEVELS OF REDUCED GLUTATHIONE IN THE
CYTOSOL. (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: NADPH + oxidized glutathione = NADP(+) + 2
glutathione.
CC -1- COFACTOR: FAD.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- MISCELLANEOUS: THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE
OXIDOREDUCTASES CLASS-I.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.emb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL; X93462; CAAG347.1; -.
 DR HSSP; P00390; 1ALG.
 DR InterPro; IPR001327; FAD_pyr_redox.
 DR InterPro; IPR000815; Hg_reductase.
 DR InterPro; IPR001100; Pyr_redox.
 DR InterPro; IPR004099; pyr_redox_2.
 DR InterPro; IPR004099; pyr_redox_dim.
 DR Pfam; PF000070; pyr_redox; 1.
 DR Pfam; PF02852; pyr_redox; 1.
 DR PRINTS; PR00368; FADPNR.
 DR PRINTS; PR00945; HGRDPRASE.
 DR PRINTS; PR00411; PNDPRASEIT.
 DR PRINTS; PR00469; PNDPRASEIT.
 DR PRODOM; PD000139; FAD_pyr_redox; 2.
 DR PROSITE; PS00076; PYRIDINE_REDOX_1; 1.
 DR Redox-active center; Oxidoreductase; Flavoprotein; FAD; NADP.
 DR INT MET 0
 FT NP BIND 3 33 FAD (ADP PART) (POTENTIAL).
 FT DISULFID 39 44 REDOX-ACTIVE (BY SIMILARITY).
 FT NP BIND 301 311 FAD (FLAVIN PART) (BY SIMILARITY).
 FT ACT_SITE 484 484 BY SIMILARITY.
 SQ SEQUENCE 499 AA; 56430 MW; 6B229901ECA0B5F7 CRC64;

Query Match 8.8%; Score 81; DB 1; Length 499;
 Best Local Similarity 24.5%; Pred. No. 5.3;
 Matches 34; Conservative 25; Mismatches 64; Indels 16; Gaps 7;

QY 29 TTEIQESEGGADEVQFAPANALPARSEAAVQPIGISOVR-----NNSKEKDLGT 84
 DB 123 TKDNNKNDQENLNEE--ILEGRNITLAVGNKPPVKGIENTISSDEFNKESKIGI 180
 QY 85 LGY-VLGITMNVITIALGAGITLGYSYKRGKDLKEQHDQ--KYCEREMQRTLPISAPT 140
 DB 181 VGSYIAVELINIVIKRIG--ISYIFARGNRILIRKEDSEVINVLNDMKNNINIVTFA 237

QY 141 NPTCEI--VDEKTVVWATS 157
 DB 238 D-VVEIKKVDKNIHLHS 255

RESULT 4
 JAG2 HUMAN
 ID JAG2 HUMAN STANDARD; PRT: 1238 AA.
 AC 09Y219; O9UE99; Q9Y6Q0; Q9Y6P9; Q9UNK8; Q9UE17;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Jagged 2 precursor (Jagged2) (HJ2).
 GN JAG2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN 1)
 RP SEQUENCE FROM N.A. (LONG ISOFORM).
 RX MEDLINE=97459705; PubMed=9315665;
 RA Luo B., Aster J.C., Hasseljian R.P., Kuo F., Sklar J.;
 RT "Isolation and functional analysis of a cDNA for human Jagged2, a gene
 encoding a ligand for the Notch1 receptor.";
 RL Mol. Cell. Biol. 17:6057-6067(1997).
 RL 12)
 RP SEQUENCE FROM N.A. (LONG ISOFORM).
 RC TISSUE=Fetal brain;
 RX MEDLINE=99180765; PubMed=10079256;
 RA Banks G.B., Mann R.S., Miltiadis E., Henriques D., Carcangiu M.-L.,
 RA Gray A., Leiman J., Ward D., Ish-Horowitz D., Aravanis-Tsakonas S.;
 RT "Human ligands of the Notch receptor.";

RL Am. J. Pathol. 154:785-794(1999).
 RN (3)
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RC TISSUE=Bone marrow;
 RX MEDLINE=20130121; PubMed=10662552;
 RA Deng Y., Madan A., Banta A.B., Friedman C., Trask B.J., Hood L.,
 RA Li L.;
 RT "Characterization, chromosomal localization, and the complete 30-kb
 RL DNA sequence of the human Jagged2 (JAG2) gene.";
 RN Genomics 63:133-138(2000).
 RN (4)
 RP SEQUENCE OF 17-1238 FROM N.A. (LONG ISOFORM).
 RC TISSUE=Heart;
 RX MEDLINE=98145947; PubMed=9486542;
 RA Valasechi C., Ghezzi C., Ballabio A., Rugarli E.I.;
 RT "JAGGED2: a putative Notch ligand expressed in the apical ectodermal
 RL ridge and in sites of epithelial-mesenchymal interactions.";
 RL Mech. Dev. 69:203-207(1997).
 CC -1- FUNCTION: PUTATIVE NOTCH LIGAND INVOLVED IN THE MEDIATION OF NOTCH
 CC SIGNALING. INVOLVED IN LIMB DEVELOPMENT (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM/HJAG2-DEL-B6; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, PLACENTA AND SKELETAL
 CC MUSCLE AND TO A LESSER EXTENT IN PANCREAS. VERY LOW EXPRESSION IN
 CC BRAIN, LUNG, LIVER AND KIDNEY.
 CC -1- DISEASE: MAY BE ASSOCIATED TO USHER SYNDROME TYPE 1A (USH1A)
 CC WHICH DESCRIBES A CONGENITAL SENSOR DEAFNESS ASSOCIATED WITH
 CC RETINITIS PIGMENTOSA AND FEEBLE-MINDEDNESS.
 CC -1- SIMILARITY: CONTAINS 16 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 DSL DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.emb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL; AF020201; AAB71189.1; -.
 CC EMBL; AF003521; AAB61285.1; -.
 CC EMBL; AF029778; AAB84215.1; -.
 CC EMBL; AF029779; AAB84216.1; -.
 CC EMBL; AF111170; AAD15562.1; -.
 CC EMBL; Y14330; CAAT7406.1; -.
 CC HSSP; P00743; ICGF.
 CC GeneW; HGNC:6189; JAG2.
 DR MIM; 602570; -.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR001774; DSL.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF00008; EGF_14.
 DR Pfam; PF01414; DSL; 1.
 DR PRINTS; PR00010; EGFRLCOO.
 DR PRINTS; PR00011; EGFRLAMININ.
 DR SMART; SM00051; DSL; 1.
 DR SMART; SM00179; EGF_CA; 9.
 DR SMART; SM00001; EGF_1like; 6.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; 10.
 DR PROSITE; PS00022; EGF_1; 16.
 DR PROSITE; PS01186; EGF_2; 12.
 DR PROSITE; PS01187; EGF_CA; 7.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Developmental protein;
 KW Repeat; Signal; Transmembrane; Alternative splicing.
 FT SIGNAL 1 26
 FT CHAIN 27 1238 JAGGED 2.
 FT DOMAIN 27 1080 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1081 1105 POTENTIAL.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87115859; PubMed=3027580;
 RA Radeke M.J., Misko T.P., Hsu C., Herzenberg L.A., Shooter E.M.;
 RT "Gene transfer and molecular cloning of the rat nerve growth factor
 receptor.";
 RL Nature 325:593-597(1987).
 RN [2]
 RP SEQUENCE OF 1-22 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=93077038; PubMed=1446821;
 RA Metzis M., Timusk T., Allikmets R., Saarma M., Persson H.;
 RT "Regulatory elements and transcriptional regulation by testosterone
 and retinoic acid of the rat nerve growth factor receptor promoter.";
 RL Gene 121:247-254(1992).
 RN [3]
 RP STRUCTURE BY NMR OF 334-418.
 RX MEDLINE=97449145; PubMed=9305641;
 RA Liepinsh E., Ilag L.L., Otting G., Ibanez C.F.;
 RT "NMR structure of the death domain of the p75 neurotrophin receptor.";
 RL EMBJ J. 16:4999-5005(1997).
 CC -1- FUNCTION: Low affinity receptor which can bind to NGF, BDNF, NT-3,
 and NT-4. Can mediate cell survival as well as cell death of
 neural cells.
 CC -1- SUBUNIT: Homodimer; disulfide-linked. Interacts with p75NTR-
 associated cell death executor.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- PTM: N- AND O-glycosylated.
 CC -1- PTM: Phosphorylated on serine residues.
 CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
 CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as there is no content in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 or send an email to license@isb-sdb.ch).
 CC -----
 CC EMBL; X05137; CAA28783.1; -;
 DR EMBL; X61269; -; NOT_ANNOTATED_CDS.
 DR PIR; A26431; A26431.
 DR PDB; INGR; 29-JUL-97.
 DR InterPro; IPR004488; Death.
 DR InterPro; IPR001368; TNFR_C6.
 DR Pfam; PF00020; TNFR_C6; 4.
 DR Pfam; PF00531; death; 1.
 DR SMART; SMO0005; DEATH; 1.
 DR SMART; SMO0208; TNFR; 3.
 DR PROSITE; PS00652; TNFR_NGFR_1; 3.
 DR PROSITE; PS00650; TNFR_NGFR_2; 4.
 DR PROSITE; PS50017; DEATH_DOMAIN; 1.
 KW Receptor; Apoptosis; Neurogenesis; Transmembrane; Glycoprotein;
 KW Repeat; Phosphorylation; Signal; 3D-structure.
 FT SIGNAL 1 29
 FT CHAIN 30 425
 FT FT 30 425 TUMOR NECROSIS FACTOR RECEPTOR
 FT FT 30 425 SUPERFAMILY MEMBER 16.
 FT DOMAIN 30 251
 FT TRANSMEM 252 273
 FT DOMAIN 274 425
 FT REPEAT 32 65
 FT REPEAT 67 108
 FT REPEAT 109 147
 FT REPEAT 149 189
 FT DOMAIN 354 419
 FT DOMAIN 198 249
 FT DISULFID 33 44
 FT DISULFID 45 58
 FT DISULFID 48 65
 FT BY SIMILARITY.
 FT BY SIMILARITY.
 FT BY SIMILARITY.

FT DISULFID 68 84
 FT DISULFID 87 100
 FT DISULFID 90 108
 FT DISULFID 110 123
 FT DISULFID 126 139
 FT DISULFID 129 147
 FT DISULFID 150 165
 FT DISULFID 168 181
 FT DISULFID 171 189
 FT CARBOHYD 61 61
 FT CARBOHYD 71 71
 SQ SEQUENCE 425 AA; 45432 MW; B2E152D94D3827F8 CRC64;
 Query Match 8.4%; Score 76.5; DB 1; Length 425;
 Best Local Similarity 24.4%; Pred. No. 12;
 Matches 42; Conservative 23; Mismatches 90; Indels 17; Gaps 5;
 QY 3 EAGVPEKRPCEDLACPEFTSQALPAFTTEIQEASSEGGADEVQVAPANALPARSEAAV 62
 Db 175 EQRLRECTPMADACEIEFGKWIPIRSTPPGSDSTABSTOEPEV-PPEQDLVPSTVADWV 233
 QY 63 QPVIGISQRYRMNSKEKDKLTGLGVYIGITMVIYIAIGAGIIGVYKRGKDLKEQHDQ 122
 Db 234 TTVWGSSQPVVTR-----GTTDNLIPV-YCSILAAVVGVLAIFYAFKRMNSCK-QNKQ 284
 QY 123 KVCEREMQRITLPISATNPFCEIVDEKTVVHTSQPVDPQSGSTPLMGQA 174
 Db 285 GANSR-----PVGQPPPEGEKTHSDSGISVDSQSLHDQQTHTQASQA 329
 RESULT 7
 ID GYG2 HUMAN STANDARD; PRT; 501 AA.
 AC 015486; 015489; 015490; 015486; 015485; 015487;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Glycogenin-2 (EC 2.4.1.186) (GN2).
 GN GYG2.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC TISSUE=Liver;
 RX MEDLINE=98010589; PubMed=9346895;
 RA Mu J., Skurat A.V., Roach P.J.;
 RT "Glycogenin-2, a novel self-glucosylating protein involved in liver
 glycogen biosynthesis.";
 RL J. Biol. Chem. 272:27589-27597(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20184741; PubMed=10721716;
 RA Zhai L., Mu J., Zong H., DePaoli-Roach A.A., Roach P.J.;
 RT "Structure and chromosomal localization of the human glycogenin-2 gene
 GYG2.";
 RL Gene 242:229-235(2000).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM BETA).
 RC TISSUE=Skin;
 RA Strausberg R.;
 RN Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
 RP CHARACTERIZATION, AND CARBOHYDRATE-LINKAGE SITE TYR-228.
 RX MEDLINE=99074257; PubMed=9857012;
 RA Mu J., Roach P.J.;
 RT "Characterization of human glycogenin-2, a self-glucosylating
 initiator of liver glycogen metabolism.";
 RL J. Biol. Chem. 273:34850-34856(1998).
 RN [5]
 RP MASS SPECTROMETRY.
 RC TISSUE=Breast cancer;

| | | | |
|----|--|------------------------------------|--|
| RX | | MEDLINE=21829517; PubMed=11840567; | |
| RA | Harris R.A., Yang A., Stein R.C., Lucy K., Brusten L., Hereth A., | | |
| RA | Parekh R., Waterfield M.D., O'Hare M.J., Neville M.A., Page M.J., | | |
| RA | Zvelebil M.J.; | | |
| RT | "Cluster analysis of an extensive human breast cancer cell line | | |
| RL | protein expression map database." | | |
| CC | - FUNCTION: SELF-GLUCOSYLATES, VIA AN INTER-SUBUNIT MECHANISM, TO | | |
| CC | FORM AN OLIGOSACCHARIDE PRIMER THAT SERVES AS SUBSTRATE FOR | | |
| CC | GLYCOGEN SYNTHASE. | | |
| CC | - CATALYTIC ACTIVITY: UDP-glucose + glycogenin = UDP + | | |
| CC | glycogylglycogenin. | | |
| CC | - COFACTOR: SELF-GLUCOSYLATION IS DEPENDENT ON THE PRESENCE OF | | |
| CC | DIVALENT METAL IONS OF WHICH MANGANESE ION IS THE MOST EFFECTIVE. | | |
| CC | - PATHWAY: GLYCOGEN BIOSYNTHESIS. | | |
| CC | - SUBUNIT: HOMODIMER, TIGHTLY COMPLEXED TO GLYCOGEN SYNTHASE. | | |
| CC | - ALTERNATIVE PRODUCTS: AT LEAST 6 ISOFORMS, ALPHA (SHOWN HERE), | | |
| CC | BETA, GAMMA, DELTA, EPSILON AND ZETA; ARE PRODUCED BY ALTERNATIVE | | |
| CC | SPLICING. | | |
| CC | - TISSUE SPECIFICITY: EXPRESSED PREFERENTIALLY IN LIVER, HEART, AND | | |
| CC | PANCREAS. | | |
| CC | - PMV: SELF-GLUCOSYLATED BY THE TRANSFER OF GLUCOSE RESIDUES FROM | | |
| CC | UDP-GLUCOSE TO ITSELF, FORMING AN ALPHA-1,4-GLYCAN OF AROUND 10 | | |
| CC | RESIDUES ATTACHED TO TYR-228. | | |
| CC | - MASS SPECTROMETRY: MW=55211.89; METHOD=MALDI. | | |
| CC | - SIMILARITY: BELONGS TO THE GLYCOGENIN FAMILY. | | |
| CC | ----- | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | |
| CC | use by non-profit institutions as long as its content is in no way | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | |
| CC | or send an email to license@isb-sib.ch). | | |
| CC | ----- | | |
| DR | EMBL; U94362; AAB84377.1; - | | |
| DR | EMBL; U94363; AAB84378.1; - | | |
| DR | EMBL; U94364; AAB84379.1; - | | |
| DR | EMBL; U94365; AAB84379.1; - | | |
| DR | EMBL; U94357; AAB84373.1; - | | |
| DR | EMBL; U94358; AAB84374.1; - | | |
| DR | EMBL; U94360; AAB84375.1; - | | |
| DR | EMBL; U94361; AAB84376.1; - | | |
| DR | EMBL; AF179624; AAF61855.1; - | | |
| DR | EMBL; AF179615; AAF61855.1; JOINED. | | |
| DR | EMBL; AF179616; AAF61855.1; JOINED. | | |
| DR | EMBL; AF179617; AAF61855.1; JOINED. | | |
| DR | EMBL; AF179618; AAF61855.1; JOINED. | | |
| DR | EMBL; AF179619; AAF61855.1; JOINED. | | |
| DR | EMBL; AF179620; AAF61855.1; JOINED. | | |
| DR | EMBL; AF179621; AAF61855.1; JOINED. | | |
| DR | EMBL; AF179622; AAF61855.1; JOINED. | | |
| DR | EMBL; AF179623; AAF61855.1; JOINED. | | |
| DR | EMBL; BC023152; AAH23152.1; - | | |
| DR | GeneW; HGNC:4700; GYG2. | | |
| DR | MTM; 300198; - | | |
| DR | InterPro: IPRO02495; GT 8. | | |
| DR | Pfam; PF01501; Glyco_transf_8; 1. | | |
| KW | transferase; Glycogen biosynthesis; Alternative splicing; | | |
| KW | Glycoprotein; Polymorphism. | | |
| FT | ACT SITE | 119 | |
| FT | CARBOHYD | 228 | |
| FT | VARSPLC | 3 | |
| FT | VARSPLC | 33 | |
| FT | VARSPLC | 42 | |
| FT | VARSPLC | 378 | |
| FT | VARSPLC | 407 | |
| FT | VARSPLC | 501 | |
| FT | VARSPLC | 413 | |
| FT | VARIANT | 270 | |
| FT | MTAGN | 228 | |
| FT | MTAGN | 230 | |
| FT | CONFLICT | 313 | |
| FT | CONFLICT | 413 | |
| FT | CONFLICT | 462 | |
| FT | CONFLICT | 464 | |

| | | | | |
|----|--|---------|-----------------------------|---------------------------------------|
| SQ | SEQUENCE | 501 AA; | 55211 MW; | 2EDED5FDAD5A7657 CRC64; |
| | Query Match | | 8.4%; | Score 76.5; DB 1; Length 501; |
| | Best Local Similarity | 27.3%; | Pred. No. 14; | |
| | Matches | 27; | Conservative | 16; Mismatches 31; Indels 25; Gaps 5; |
| OY | 6 VPEKPCED-IRCEETTSQAL-----DAFTTE-----IQASSEGDADEVQFA | 48 | | |
| | : : : : : : : : : : : : : : : : | | | |
| DB | 369 LPEGRRSDMDIACPETEPFATVTCDPISQPSPGADAFETETILQIPANKVSVSSEETFE | 428 | | |
| | : : : : : : : : : : : : : : : : | | | |
| OY | 49 PANALPARS-----EAAAVQPIGISQRNRNMSKEKX | 80 | | |
| | : : : : : : : : : : : : : : : : | | | |
| DB | 429 PSQLPALRLDRPSLQDALEVLAVSVSQ-ISIEEKVAE | 466 | | |
| | : : : : : : : : : : : : : : : : | | | |
| | RESULT 8 | | | |
| ID | SPXD_SYNY3 STANDARD; PRT; | 505 AA. | | |
| AC | PS54735; | | | |
| DT | 01-OCT-1996 (Rel. 34, Created) | | | |
| DD | 01-OCT-1996 (Rel. 34, Last sequence update) | | | |
| DT | 16-OCT-2001 (Rel. 40, Last annotation update) | | | |
| DE | Probable serine/threonine-protein kinase D (EC 2.7.1.-). | | | |
| GN | SPKD OR SLI0776. | | | |
| OS | Synechocystis sp. (strain PCC 6803). | | | |
| OC | Bacteria; Cyanobacteria; Chroococcales; Synechocystis. | | | |
| OX | NCBI_TaxID=1148; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | Kamel A., Ikeuchi M.; | | | |
| RT | "A novel gene, <i>spkD</i> , encodes active Ser/Thr protein kinase in the | | | |
| RL | motile cyanobacterium <i>Synechocystis</i> sp. PCC 6803." | | | |
| RL | Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases. | | | |
| RP | [2] | | | |
| RX | SEQUENCE FROM N.A. | | | |
| RY | MEDLINE=96127529; PubMed=85590279; | | | |
| RA | Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N., | | | |
| RA | Sugitara T., Tabata S.; | | | |
| RT | "Sequence analysis of the genome of the unicellular cyanobacterium | | | |
| RT | <i>Synechocystis</i> sp. strain PCC6803. I. Sequence features in the 1 Mb | | | |
| RL | region from map positions 64% to 92% of the genome."; | | | |
| RL | DNA Res. 2:153-166(1995). | | | |
| CC | - SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. | | | |
| CC | ----- | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | | |
| CC | or send an email to license@isb-sib.ch). | | | |
| CC | ----- | | | |
| DR | EMBL; AB046600; BAB17036.1; - | | | |
| DR | EMBL; D64005; BAA10726.1; - | | | |
| DR | InterPro; IPR000719; Euk_pkinase. | | | |
| DR | InterPro; IPR003646; SH3_bac. | | | |
| DR | InterPro; IPR002290; Ser_thr_kinase. | | | |
| DR | Pfam; PF00069; pkinase; I. | | | |
| DR | ProDom; PD000001; Euk_pkinase; 1. | | | |
| DR | SMART; SM00287; SH3b; 1. | | | |
| DR | PROSITE; PS00107; PROTEIN_KINASE_ATP; 1. | | | |
| DR | PROSITE; PS00108; PROTEIN_KINASE_ST; 1. | | | |
| DR | PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1. | | | |
| KW | Hypothetical protein; transferase; Serine/threonine-protein kinase; | | | |
| KW | ATP-binding; Complete proteome. | | | |
| FT | DOMAIN 9 .. 271 PROTEIN KINASE. | | | |
| FT | NP BIND 15 .. 23 ATP (BY SIMILARITY). | | | |
| FT | BT BINDING 40 .. 40 ATP (BY SIMILARITY). | | | |
| FT | ACT SITE 136 .. 136 BY SIMILARITY. | | | |
| SQ | SEQUENCE 505 AA; 55213 MW; C4FI2A1886C4D51C CRC64; | | | |
| | Query Match | 8.3%; | Score 76; DB 1; Length 505; | |
| | Best Local Similarity | 27.1%; | Pred. No. 16; | |

Matches 35; Conservative 17; Mismatches 69; Indels 8; Gaps 4;

QY 8 EKRCEDLRCEPTSQLPATTTCIOASGPGADGVQVAPANA-IPARSEAAVQPV- 65
DB 250 EMESRRYSSAANYQALHSL--ISSGAPALPMETVRAPSPNEFLVTRSTKAEVTV 306
QY 66 --IGISORVNMNSKEKDLGTLGYVLGTMVVIITAGAGIIGVSKRGDKJKEQHDQK 123
DB 307 KPVGNSHNNNSNNNGSKINTLTLV-IGIIVTAGLGSGFITTOQIKENARAQAQKE 365
QY 124 VCEREMQRI 132
DB 366 KOEAEQKRI 374

RESULT 9
TNR8_HUMAN
ID TNR8_HUMAN STANDARD; PRT; 595 AA.
AC P28908;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 8 precursor (CD30L receptor) [lymphocyte activation antigen CD30] (KI-1 antigen).
GN TNFRSF8 OR CD30.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (LONG ISOFORM).
RC TISSUE=Lymphoid;
RX MEDLINE=92154659; PubMed=1310894;
RA Duerkop H., Latza U., Hummel M., Eitelbach F., Seed B., Stein H.;
RT "Molecular cloning and expression of a new member of the nerve growth factor receptor family that is characteristic for Hodgkin's disease";
RT Cell 68:421-427(1992).
RL [2]
RN SEQUENCE FROM N.A. (LONG ISOFORM).
RP MEDLINE=95089787; PubMed=7527901;
RA Jung W., Krueger S., Renner C., Gause A., Sahin U., Trumper L.,
RA Pfeundschuh M.;
RT "Opposite effects of the CD30 ligand are not due to CD30 mutations: results from cDNA cloning and sequence comparison of the CD30 antigen from different sources";
RT Mol. Immunol. 31:1329-1334(1994).
RL [3]
RN SEQUENCE FROM N.A. (SHORT ISOFORM).
RP MEDLINE=96437016; PubMed=8839832;
RA Horie R., Ito K., Tatewaki M., Nagai M., Alzawa S.,
RA Higashihara M., Ishida T., Inoue J., Takizawa H., Watanabe T.;
RT "A variant CD30 protein lacking extracellular and transmembrane domain is induced in HL-60 by tetradecanoylphorbol acetate and is expressed in alveolar macrophages";
RT Blood 88:2422-2432(1996).
RL [4]
RN -1- FUNCTION: Receptor for TNFRSF8/CD30L. May play a role in the regulation of cellular growth and transformation of activated lymphoblasts. Regulates gene expression through activation of NF-kappaB.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (long isoform); Cytoplasmic (short isoform).
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a short form/cytoplasmic/CD30V; are produced by alternative initiation.
CC -1- PTM: PHOSPHORYLATED ON SERINE AND TYROSINE RESIDUES.
CC -1- DISEASE: MOST SPECIFIC HODGKIN'S DISEASE ASSOCIATED ANTIGEN.
CC -1- SIMILARITY: CONTAINS 6 TNR- CYS REPEATS.
CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD30 entry;
CC WWW=ftp://www.ncbi.nlm.nih.gov/prov/cd/cd30.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcement/> or send an email to license@ebi.ac.uk).

CC EMBL; M83554; AAA51947.1; -
CC EMBL; S75768; AAD14188.1; -
CC EMBL; D86042; BAI12973.1; -
CC PIR; A42086; A42086;
CC GeneW; HGNC:11923; TNFRSF8.
CC MIM; 153243; -
CC InterPro; IPR001368; TNFR_c6.
CC Pfam; PF00020; TNFR_c6; 4.
CC SMART; SM00208; TNFR; 4.
CC PROSITE; PS00652; TNFR_NGFR_1; 2.
CC PROSITE; PS50050; TNFR_NGFR_2; 2.
CC KW Receptor; Transmembrane; Glycoprotein; Signal; Repeat; Phosphorylation; Alternative initiation.
FT SIGNAL 1 18
FT CHAIN 19 595
FT FT
FT CHAIN 464 595
FT INT MET 464 464
FT DOMAIN 19 379
FT TRANSREM 380 407
FT DOMAIN 408 595
FT REPEAT 28 66
FT REPEAT 68 106
FT REPEAT 107 150
FT REPEAT 205 241
FT REPEAT 243 281
FT REPEAT 282 325
FT DOMAIN 347 377
FT DISULFID 29 44
FT DISULFID 45 58
FT DISULFID 48 65
FT DISULFID 69 81
FT DISULFID 84 98
FT DISULFID 87 106
FT DISULFID 108 122
FT DISULFID 131 149
FT DISULFID 233 240
FT DISULFID 244 256
FT DISULFID 259 273
FT DISULFID 262 281
FT DISULFID 283 297
FT DISULFID 289 300
FT CARBOHYD 101 101
SQ SEQUENCE 595 AA; 63747 MW; 7A407CC78A6E0B8C CRC64;
Query Match 8.2%; Score 75; DB 1; Length 595;
Best Local Similarity 23.1%; Pred. No. 23;
Matches 43; Conservative 31; Mismatches 70; Indels 42; Gaps 10;

QY 13 EDRCETTSQALPAT---TEIQASGPGADGVQVAPANA-IPARSEAAVQPVIGI 68
DB 315 QDAEKDTTFEAPLGTGPDNCNTPENGEPAST-----SPQSLIVDSQASKTLP-IPT 368
QY 69 SQRRVNMNSKEKDLGTLGYVLGTMVVIITAGAGIIGVSKRGDKJKEQHDQVCEKE 128
DB 369 SAAVVALSTGCKPVLD-GLVLFVILVIVVVGSSAFL-----LCRRACKRR 415
QY 129 MOR---TLPLSAFTNPTCEIVDEKTVVHTSQT-----PVDDQEG--STPLMGQ 173
DB 416 IRQKHLICYPVOT-SQPKLELVDSRP---RRSTQLRSGASVTEVVAERGLMGSPLMET 471
QY 174 AGPGGA 179
DB 472 CHSVGA 477


```
RESULT 10
YFGF_ECOLI
ID YFGF_ECOLI STANDARD; PRT; 747 AA.
AC P77172;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yfgf.
GN YFGF OR B2503.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
GN NCBI_TaxID=562;
OX
RN
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RL "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN
RP SEQUENCE FROM N.A.
RX STRAIN=K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsunishi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimura H.,
RA Ohshima T., Oyama S., Saito N., Sempel G., Satoh Y., Sivasubaram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horinouchi T.;
RL "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT K-12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features.";
RL DNA Res. 4:91-113(1997).
CC -1- SIMILARITY: CONTAINS 1 DUF2 DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
CC
CC EMBL; AE000336; AAC75556.1; -.
DR EMBL; D90880; BAA16393.1; -.
DR ECGene; EG14202; YFGF.
DR InterPro; IPR001633; EAL.
DR InterPro; IPR001601; GCDER.
DR Pfam; PF00563; EAL; 1.
DR SMART; SM00267; DUF1; 1.
DR SMART; SM00052; DUF2; 1.
KW Hypothetical protein; Complete proteome.
FT DOMAIN 492 736 DUF2.
SQ SEQUENCE 747 AA; 85607 MW; 7D5EBB0E646C8EDF CRC64;

Query Match 8.1%; Score 74.5; DB 1; Length 747;
Best Local Similarity 25.2%; Pred. No. 33;
Matches 31; Conservative 15; Mismatches 52; Indels 25; Gaps 3;

QY 45 QVFAANLIPARSEAAVQPIVIGISQVRNMSKEKDGTL-----GVVGIITMMVI 96
DB 141 QPAAVGLASRENLVGVPF-----NIGTLINYOALLVGNLIGVPLCYF 185
QY 97 IIAIGAGIILIGSYRGKDLKEOHDKVCERBMQRITLPLSAFTNPCEIVDEKTVVHT 156
DB 186 IIRVRNPFYLRASY--SOLKQVDAAKVTKEFALMLALGALLLLCMPLNEKSTFST 243
QY 157 SQT 159
```

```
Db 244 NYT 246

RESULT 11
MTKB_METEX
ID MTKB_METEX STANDARD; PRT; 260 AA.
AC P53595;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Malate--CoA ligase alpha chain (EC 6.2.1.9) (Malyl-CoA synthetase)
DE (Malate thiokinase) (MTK-alpha).
GN MTKB.
OS Methylobacterium extorquens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_TaxID=408;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=AM1 / NCIMB 9133;
RX MEDLINE=95050329; PubMed=7961516;
RA Chistoserdova L.V., Lidstrom M.E.;
RL "Genetics of the serine cycle in Methylobacterium extorquens AM1:
RT identification, sequence, and mutation of three new genes involved in
RT C1 assimilation, orf4, mtkA, and mtkB.";
RL J. Bacteriol. 176:7398-7404(1994).
CC -1- CATALYTIC ACTIVITY: ATP + malate + CoA = ADP + phosphate + malyl-
CC CoA.
CC -1- PATHWAY: SERINE CYCLE.
CC -1- SUBUNIT: COMPOSED OF AN ALPHA CHAIN AND A BETA CHAIN.
CC -1- SIMILARITY: TO OTHER ALPHA SUBUNITS OF SUCCINYL-COA SYNTHETASE,
CC OF MALATE--COA LIGASE AND TO ATP CITRATE-LYASE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
CC
CC EMBL; L33465; AAA62655.1; -.
DR HSP; P07459; 28CU.
DR InterPro; IPR003781; CoA binding.
DR InterPro; IPR003033; CoA ligase.
DR Pfam; PF00549; ligase-CoA; 1.
DR Pfam; PF02829; CoA binding; 1.
DR PROSITE; PS01216; SUCCINYL-CoA_LIG_1; 1.
DR PROSITE; PS00399; SUCCINYL-CoA_LIG_2; 1.
KW Ligase; Phosphorylation; ATP-binding.
FT ACT_SITE 251 251 PHOSPHORYLATED IN THE COURSE OF
FT CATALYSIS (BY SIMILARITY).
SQ SEQUENCE 260 AA; 26978 MW; 9A26A3B057A4A447 CRC64;

Query Match 8.0%; Score 73; DB 1; Length 260;
Best Local Similarity 22.6%; Pred. No. 14;
Matches 43; Conservative 30; Mismatches 65; Indels 52; Gaps 8;

QY 21 TSOALPAFTTIOESSEGA--DEVQVAPAPNALPARESEA-----AYQVY 65
DB 47 THGVGVFNT-VKEAVEATGATSTTFVAPPPAADALMEADAGLKVCSITDGIAPADM 105
QY 66 IGISQVRNMSKEKDLGTLGYVLGITMMV-----IIAIG--AGIILIGSYRGKDLK 117
DB 106 MRVKRYLRKRYPEKR-----TWVGNPCAGIISPGKSMGIMGHILYLGK--- 151
QY 118 EQHDKVCER-----EMQRITLPLSAFTNPCEIVDEKTVVHTSQTVPVDPQSG 167
DB 152 ---VGVISRGTLGYEAQAOMKEIGIGISTGVIGGDPINGSSFLDHALPEQDPTEA 207
QY 168 TPLMGQAGTP 177
```

Db 208 VLMIGIGP 217

RESULT 12
IDENT_SCHPO STANDARD; PRT; 536 AA.

AC 013655;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical serine-rich protein C11G7.01 in chromosome I.
GN SPAC11G7.01.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetaceae; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
NCBI_TaxID=4896;
RN NCB1
RP
SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21648401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Baeham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean U.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,
RA Welfjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs W., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motier S.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motier S.,
RA Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez S., del Rey F., Benito J.,
RA Dominguez A., Revelante J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; 299161; CAB16206.1; -
KW Hypothetical protein; Transmembrane.
FT DOMAIN 25 238 SRR-RICH.
FT TRANSMEM 247 267 POTENTIAL.
FT TRANSMEM 351 371 POTENTIAL.
SQ SEQUENCE 536 AA; 52992 MW; CA7EFAF19ED658CA CRC64;

Query Match 8.0%; Score 73; DB 1; Length 536;
Best Local Similarity 36.4%; Pred. No. 31;
Matches 20; Conservative 10; Mismatches 21; Indels 4; Gaps 1;

Qy 76 SKRKDLGLGYLG---ITMWTIIAIGAGIIGSYKRGKDKQEQHDQKVC 126
Db 232 SKSKNSNHWGVGCVAVIPGVVILILGLGIFLMKHKQKRIKARMEVEE 286

RESULT 13

ICP0_HSVBU
ID_ICP0_HSVBU STANDARD; PRT; 676 AA.

AC P29128;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Trans-acting transcriptional protein ICP0 (p135 protein) (IER
DE 2.9/ER.6).
GN ICP0.
OS Bovine herpesvirus type 1 (strain Jura).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirinae.
NCBI_TaxID=31518;
RN NCB1
RP
SEQUENCE FROM N.A.
RX MEDLINE=92219360; PubMed=1313901;
RA Wirth U.V., Fraefel C., Vogt E., Vlcek C., Paces V., Schwyzler M.,
RT "Immediate-early RNA 2.9 and early RNA 2.6 of bovine herpesvirus 1
RT are 3' coterminal and encode a putative zinc finger transactivator
RT protein.";
RL J. Virol. 66:2763-2772(1992).

CC -1- PTM: THE STRONGLY ACIDIC REGION MIGHT SERVE AS A TRANSCRIPTIONAL
CC ACTIVATION DOMAIN, POSSIBLY REGULATED THROUGH PHOSPHORYLATION BY
CC CASEIN KINASE II.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -1- SIMILARITY: TO OTHER HERPESVIRUSES ICP0 PROTEIN.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; M84465; AAA46062.1; -
DR PIR; A0004801; CA006138.1; -
DR PIR; B38209; EDBE23.
DR HSP; P28990; ICHC.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00097; Zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PSS0089; ZF_RING_2; 1.
KW Transcription regulation; Trans-acting factor; Activator; Zinc-finger;
KW DNA-binding; Early protein; Repressor; Phosphorylation.
FT ZN_FING 13 52 RING-TYPE.
FT DOMAIN 284 331 ASF/GHU-RICH (ACIDIC).
SQ SEQUENCE 676 AA; 67879 MW; 11B06BA4E5C4EB71 CRC64;

Query Match 8.0%; Score 73; DB 1; Length 676;
Best Local Similarity 23.2%; Pred. No. 40;
Matches 47; Conservative 21; Mismatches 95; Indels 40; Gaps 8;

Qy 2 GEAGVPEKRCEDLRCPETTSOALPFTTBIOEASGPGADVEQVAPANA----- 52
Db 336 GAGSAARARPARERQYVSTGRTPAVQAPRSLARPCRAAAVAPSSNRSGRRDP 395

Qy 53 -LPRSEAAAVPVYIGISQVRNMSKSKDLGT---LGYVLGITMWTIIAIGAGIILGY 108
Db 396 RLPAAPPAAD-----AAQARACSPREPRGAGIGVAAGET-----AGWGAGEBR 443

Qy 109 SYKRGDKLQKHQKVCER-----EMQRTLPISA---FTNPCEIVDEKTVVHTSOTP 160
Db 444 GERARLIGAGPPRVQARRRRTELDRAPTAPAPAPAPAPAPAPAPAPAPAPAP 500

Qy 161 VDPQEGSTP---LWGQAGPGCA 179
Db 501 ADAPAPAAPGPAPAGAGIGTPAA 523

RESULT 14
EFG_DEIRA

```

ID   EFG_DEIRA          STANDARD;          PRT;          698 AA.
AC   O9RKS;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DE   Elongation factor G (EF-G).
GN   FUS4 OR DR0307.
OS   Deinococcus radiodurans.
OC   Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OX   NCBI_TaxID=1299;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R1.
RX   MEDLINE=20036896; PubMed=10567266;
RA   White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA   Dodson R.J., Haft D.H., Gwin M.L., Nelson W.C., Richardson D.L.,
RA   Wolfat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA   Vamathevan J.J., Lam P., McDonald L., Ueberback T., Zaleski C.,
RA   Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA   Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA   Fraser C.M.;
RT   "Sequence of the radioresistant bacterium Deinococcus
RT   radiodurans R1."
RL   Science 286:1571-1577(1999).
CC   -1- FUNCTION: This protein promotes the GTP-dependent translocation of
CC   the nascent protein chain from the A-site to the P-site of the
CC   ribosome.
CC   -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC   EF-G/EF-2 SUBFAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; AE001891; AAF09887.1; -.
DR   HSSP; P13551; 1FNM.
DR   TIGR; DR0307; -.
DR   InterPro; IPR000640; EF-G.
DR   InterPro; IPR004161; EFU_D2.
DR   InterPro; IPR000795; EF_GTPbind.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00679; EFG_C_1.
DR   Pfam; PF00009; GTP_EFTU_1.
DR   Pfam; PF03144; GTP_EFTU_D2_1.
DR   PRINTS; PR00315; BLONGATNFC.
DR   TIGRPFAM; TIGR00484; EF-G; 1.
DR   TIGRPFAM; TIGR00231; small_GTP_1.
DR   PROSITE; PS00301; EFACOR_GTP; 1.
KW   Elongation factor; Protein biosynthesis; GTP-binding;
KM   Complete proteome.
FT   NP_BIND          20      27      GTP (BY SIMILARITY).
FT   NP_BIND          90      94      GTP (BY SIMILARITY).
FT   NP_BIND          144     147      GTP (BY SIMILARITY).
SQ   SEQUENCE 698 AA; 76798 MW; 92E03D5645B5F4B CRC64;

Query Match          8.0%; Score 73; DB 1; Length 698;
Best Local Similarity 24.8%; Pred. No. 42; Indels 18; Gaps 5;
Matches 40; Conservative 22; Mismatches 81;

OY   26 PAFTTEIQEAS-----EGFGADEVOYFAPANALPARSEAAVOPVIG-----ISQVRKMS 76
DB   444 PTFEVEEDDESQGTITAGGEHLRLVDRRLREYKYVANVGAPQVAYRETIKQVEVDS 503
OY   77 KEKKDLGTLGYVLTGMVYIITAGITIGSYRGKDLKEQ--HQQKCEEREMRITL 134
DB   504 KFNARSGRGQYGVHVLKRLREPLEPGAGIFENAVVGVGVPEYVIGPRAQKVEASMG--SG 561

```

```

OY   135 PLSAFTNPTCEIVDEKTVVHTSOTFVDPDQSGSTPLMQAG 175
DB   562 PMLGF-----PVVDIKVYIYDGSYHEVDSSEWAFKIASMG 597

RESULT 15
ID   B3AR_MOUSE        STANDARD;          PRT;          400 AA.
AC   P25962; O9QZ98;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Beta-3 adrenergic receptor.
GN   ADNR3 OR ADNR3R OR B3BAR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Swiss;
RX   MEDLINE=92037534; PubMed=1718744;
RA   Naimias C., Blin N., Elalouf J.M., Mattei M.-G., Strosberg A.D.,
RA   Emswiler L.J.;
RT   "Molecular characterization of the mouse beta3-adrenergic receptor:
RT   relationship with the atypical receptor of adipocytes."
RL   EMBO J. 10:3721-3727(1991).
RN   [2]
RP   REVISIONS, SEQUENCE FROM N.A.
RX   MEDLINE=93279311; PubMed=8389293;
RA   van Spronsen A., Naimias C., Krief S., Briand-Sutren M.-M.,
RA   Strosberg A.D., Emswiler L.J.;
RT   "The promoter and intron/exon structure of the human and mouse beta
RT   3-adrenergic-receptor genes."
RL   Eur. J. Biochem. 213:1117-1124(1993).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM B).
RC   STRAIN=C57BL/6; TISSUE=Brown adipose tissue;
RX   MEDLINE=9336494; PubMed=1045305;
RA   Evans B.A., Papaioannou M., Hamblton S., Summers R.J.;
RT   "Alternative splicing generates two isoforms of the beta3-adrenoceptor
RT   which are differentially expressed in mouse tissues."
RL   Br. J. Pharmacol. 127:1525-1531(1999).
RN   [4]
RP   SEQUENCE OF 378-400 FROM N.A.
RX   TISSUE=Adipose tissue;
RX   MEDLINE=93125503; PubMed=1336117;
RA   Graneman J.G., Labners K.N., Rao D.D.;
RT   "Rodent and human beta 3-adrenergic receptor genes contain an intron
RT   within the protein-coding block."
RL   Mol. Pharmacol. 42:964-970(1992).
CC   -1- FUNCTION: BETA-ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-
CC   INDUCED ACTIVATION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G
CC   PROTEINS. BETA-3 IS INVOLVED IN THE REGULATION OF LIPOLYSIS AND
CC   THERMOGENESIS.
CC   -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE
CC   PRODUCED BY ALTERNATIVE SPLICING.
CC   -1- TISSUE SPECIFICITY: WHITE AND BROWN ADIPOSE TISSUES, AND DIGESTIVE
CC   TRACT. ISOFORM B HIGHEST IN BRAIN.
CC   -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; X72862; CAA51384.1; -.
DR   EMBL; X60438; CAA42966.1; -.

```

DR EMBL: AF193027; AAF05768.1; -
DR EMBL: SS3290; AAB24836.1; -
DR PIR: S18548; S18548.
DR PIR: S12804; S12804.
DR MGD; MGI:87939; AATB3.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family; Phosphorylation; Lipoprotein; Palmitate;
KW Alternative splicing
FT DOMAIN 1 36 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 37 60 1 (POTENTIAL).
FT DOMAIN 61 69 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 70 88 2 (POTENTIAL).
FT DOMAIN 89 108 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 109 130 3 (POTENTIAL).
FT DOMAIN 131 152 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 153 175 4 (POTENTIAL).
FT DOMAIN 176 200 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 201 222 5 (POTENTIAL).
FT DOMAIN 223 289 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 290 311 6 (POTENTIAL).
FT DOMAIN 312 323 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 324 344 7 (POTENTIAL).
FT DOMAIN 345 400 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 8 8 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 107 186 BY SIMILARITY.
FT LIPID 358 358 PALMITATE (BY SIMILARITY).
FT VARSPLIC 388 400 REDGEGARPEPT -> SLLREPRRLTCLGYP (IN ISOFORM B).
SQ SEQUENCE 400 AA; 43006 MW; 374E072C01DFA32E CRC64;
Query Match 7.9%; Score 72.5; DB 1; Length 400;
Best Local Similarity 19.5%; Pred. No. 25;
Matches 39; Conservative 21; Mismatches 57; Indels 83; Gaps 8;
QY 7 PEKPCEDLCPETTSQALPAFTTEIQEASBEGADDEVQVFPANALPARSEMAAVQVPI 66
DB 244 PESSPPSPSPSPSPATGTPA-----APDGVPPCGRRPARLPL- 282
QY 67 GISQRYMNSKEKKDLGTLGYLGITM-----VITIAIGAIL-----IG 107
DB 283 -----REHRLATTLGLIMGIFSLCLMPLFLANVLRALAGPSLVPSGVFIALNMLG 332
QY 108 YS-----YKRGKDLKEQHDKVCE-----REMQRITLPLSAFTNPTCEIVDEKTV 153
DB 333 YANSAFNPVLYCHSPDFRDAFRLLCSYGGRPEBRPAVTFPASP----- 377
QY 154 VHTSQTP---VDPQEGSTP 169
DB 378 VEARQSPPLNRFQDEGARP 397

Search completed: April 7, 2003, 09:17:01
Job time : 12.1871 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:25 / Search time 48.0062 Seconds

(without alignments)
768.284 Million cell updates/sec

Title: US-10-057-951-2_COPY_85_263

Sequence: 1 SGEAGVPEKPCEDLRCPET.....PVDPGSTPLMGQAGTPGA 179

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

```
SPTREMBL_21.*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mtc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.rvins:*
16: sp.bacteriap:*
17: sp.archaeap:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|---------------------|
| 1 | 916 | 100.0 | 263 | 4 Q96FE7 | Q96FE7 homo sapien |
| 2 | 913 | 99.7 | 263 | 4 Q00318 | Q00318 homo sapien |
| 3 | 913 | 99.7 | 263 | 4 Q00318 | Q00318 homo sapien |
| 4 | 93.5 | 10.2 | 8805 | 5 Q9V6V4 | Q9V6V4 drosophila |
| 5 | 87.5 | 9.6 | 759 | 5 Q20429 | Q20429 caenorhabdi |
| 6 | 86 | 9.4 | 290 | 5 Q9BL62 | Q9BL62 caenorhabdi |
| 7 | 86 | 9.4 | 1165 | 4 Q9C0G0 | Q9C0G0 homo sapien |
| 8 | 85.5 | 9.3 | 1115 | 11 Q91V14 | Q91V14 mus musculu |
| 9 | 85.5 | 9.3 | 1116 | 11 Q63633 | Q63633 rattus norv |
| 10 | 84.5 | 9.2 | 607 | 17 Q8U3P9 | Q8U3P9 pyrococcus |
| 11 | 83.5 | 9.1 | 517 | 16 Q93J57 | Q93J57 streptomyce |
| 12 | 83.5 | 9.1 | 733 | 16 Q9PA31 | Q9PA31 xyloella fas |
| 13 | 83 | 9.1 | 1553 | 5 Q95077 | Q95077 caenorhabdi |
| 14 | 82 | 9.0 | 774 | 10 Q9L128 | Q9L128 oryza sativ |
| 15 | 81.5 | 8.9 | 535 | 10 Q8W4V9 | Q8W4V9 solanum tub |
| 16 | 81.5 | 8.9 | 715 | 10 Q9M839 | Q9M839 arabidopsis |

| | | | | | |
|----|------|-----|------|-----------|---------------------|
| 17 | 81.5 | 8.9 | 1365 | 5 Q9NEV4 | Q9NEV4 caenorhabdi |
| 18 | 80.5 | 8.8 | 803 | 16 Q25153 | Q25153 helicobacte |
| 19 | 80.5 | 8.8 | 918 | 16 Q8XR11 | Q8XR11 ralsstonia s |
| 20 | 80.5 | 8.8 | 1852 | 3 Q9C2H4 | Q9C2H4 neurospora |
| 21 | 80.5 | 8.8 | 2358 | 16 Q9L1V8 | Q9L1V8 streptomyce |
| 22 | 80 | 8.7 | 618 | 2 Q9F5E5 | Q9F5E5 agrobacteri |
| 23 | 79.5 | 8.7 | 770 | 2 Q9AQ89 | Q9AQ89 bacillus th |
| 24 | 79 | 8.6 | 960 | 16 Q9ZKP5 | Q9ZKP5 rhizobium m |
| 25 | 78.5 | 8.6 | 457 | 16 Q8UHM4 | Q8UHM4 agrobacteri |
| 26 | 78.5 | 8.6 | 478 | 5 Q25684 | Q25684 plasmodium |
| 27 | 78.5 | 8.6 | 552 | 5 Q9VE62 | Q9VE62 drosophila |
| 28 | 78.5 | 8.6 | 2972 | 5 P90891 | P90891 caenorhabdi |
| 29 | 78 | 8.5 | 224 | 16 Q8Z5M1 | Q8Z5M1 salmoneilla |
| 30 | 78 | 8.5 | 336 | 17 Q8ZDP0 | Q8ZDP0 pyrobaculum |
| 31 | 78 | 8.5 | 352 | 16 Q98BL5 | Q98BL5 rhizobium l |
| 32 | 77.5 | 8.5 | 267 | 16 Q8UB02 | Q8UB02 agrobacteri |
| 33 | 77.5 | 8.5 | 947 | 4 Q8TDV8 | Q8TDV8 homo sapien |
| 34 | 77.5 | 8.5 | 947 | 4 Q8TBZ2 | Q8TBZ2 homo sapien |
| 35 | 77.5 | 8.5 | 1230 | 5 Q20626 | Q20626 caenorhabdi |
| 36 | 77 | 8.4 | 624 | 11 Q70236 | Q70236 mus musculu |
| 37 | 76.5 | 8.4 | 518 | 3 Q9F732 | Q9F732 neurospora |
| 38 | 76.5 | 8.4 | 766 | 2 Q9ZHF3 | Q9ZHF3 neisseria m |
| 39 | 76.5 | 8.4 | 1101 | 4 Q9ULP4 | Q9ULP4 homo sapien |
| 40 | 76.5 | 8.4 | 1116 | 4 Q9H2X9 | Q9H2X9 homo sapien |
| 41 | 76.5 | 8.4 | 1116 | 4 Q9H4Z0 | Q9H4Z0 homo sapien |
| 42 | 76 | 8.3 | 453 | 16 Q8UUV0 | Q8UUV0 agrobacteri |
| 43 | 76 | 8.3 | 482 | 5 Q93998 | Q93998 plasmodium |
| 44 | 76 | 8.3 | 847 | 6 Q95KP7 | Q95KP7 canis fami1 |
| 45 | 76 | 8.3 | 1406 | 5 Q01777 | Q01777 caenorhabdi |

ALIGNMENTS

RESULT 1

ID Q96FE7 PRELIMINARY; PRT; 263 AA.

AC Q96FE7; 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Unknown (protein for MG:17330).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID:9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN;

RA Strauberg R.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC011049; AA011049.1; -

DR InterPro; IPR000001; Kringle.

DR Pfam; PF00051; Kringle; 1.

DR Prodom; PD000195; Kringle; 1.

DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.

DR PROSITE; PS00070; KRINGLE_2; 1.

SQ SEQUENCE 263 AA; 28234 MW; 197C3BEE88FA242 CRC64;

Query Match 100.0%; Score 916; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 6.6e-78;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db 205 DQKVCEREMORITLPLSAFTNPTCEIVDEKTVVHTSQTPVDPQEGSTPLMGAGTPGA 263
RESULT 2
000318 PRELIMINARY; PRT; 263 AA.
AC 000318;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE MGSCC.DJ515N1.2 protein.
GN MGSCC.DJ515N1.2.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Du Z., Scheet P., Harper M.,
RT "The sequence of H. sapiens PAC clone RP3-S15N1.",
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC002073; AAB54054.1; -.
DR HSSP; P00749; 1KDU.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE; FALSE_NEG.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00023; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
SQ SEQUENCE 263 AA; 28248 MW; 197C3EEB854A242 CRC64;

Query Match 99.7%; Score 913; DB 4; Length 263;
Best Local Similarity 99.4%; Pred. No. 1.3e-77;
Matches 178; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SGEAGVPEKRCEDLRCPEITTSQALPAFTTEIQDASGPGADVQVPAFANALPARSEAA 60
Db 85 SGEAGVPEKRCEDLRCPEITTSQALPAFTTEIQDASGPGADVQVPAFANALPARSEAA 144
Qy 61 AVQPVIGISQVRNMSKEKDLGTLGYVLGITMVAIIAIGAGIIGSYKRGKDLKEQH 120
Db 145 AVQPVIGISQVRNMSKEKDLGTLGYVLGITMVAIIAIGAGIIGSYKRGKDLKEQH 204
Qy 121 DQKVCEREMORITLPLSAFTNPTCEIVDEKTVVHTSQTPVDPQEGSTPLMGAGTPGA 179
Db 205 DQKVCEREMORITLPLSAFTNPTCEIVDEKTVVHTSQTPVDPQEGSTPLMGAGTPGA 263

RESULT 3
09V6V4 PRELIMINARY; PRT; 8805 AA.
AC 09V6V4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Shot protein.
GN SHOT OR CG18076 OR CG18637.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BERKELEY;
RX MEDLINE=20196006; Pubmed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

```

Db 3435 GTPAEG-GSNLQNVKDAKALGALA-IVGAPVL-----AGKALDALTTTEKLNREH 3489
 QY 121 --DQRYCEREMQRTPLPLSAFTNPCEIVDEKTVVHTSOTPVDPQEGSTP 169
 Db 3490 STSQVTEVEE---LPLEIITTSATLVTKVTTT---TTTTTSTSG-TP 3534

RESULT 4

Q9M261 PRELIMINARY; PRT; 569 AA.
 AC Q9M261;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE Hypothetical 63.5 kDa protein.
 GN F7M19.10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Nakamura G., Fartmann B., Dauner D., Sterr W., Holland R.,
 RA Weichselgartner M., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.,
 RA Quecter F., Salanoubat M.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL338643; CAB6471.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 569 AA; 63537 MW; B75942AC7C565F5 CRC64;

Query Match

Best Local Similarity 10.2%; Score 93.5; DB 10; Length 569;
 Matches 27; Conservative 10; Mismatches 36; Indels 11; Gaps 1;

QY 8 EKRPCEDLRCPEITTSQALPAFTTEIOASBEGADVEQVAPAN-----ALPAA 56
 Db 301 KKRAERRRVVEESGPAEPGAREVPEVAAPADVDVDPAPRPTETATQAVIALPAR 360
 QY 57 SEAAVQVPIGISQVRMNSKEKK 80
 Db 361 DKASGKSPQIDTSQEKRRKKKKK 384

RESULT 5

Q20429 PRELIMINARY; PRT; 759 AA.
 AC Q20429;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE F45B12.2 protein.
 GN F45B12.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=BRISTOL N2;
 MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kerstew J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders R., Showkhen R.,
 RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,

RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sprot J., Wohldman P.; III of C.
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Latreille P.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U29516; AAA68790.1; -.
 DR InterPro; IPR004366; Cyclin.
 DR InterPro; IPR000812; TPPIB euk.
 DR Pfam; PF00382; transcript_fac2; 2.
 DR SMART; SM00385; CYCLIN_2.
 SQ SEQUENCE 759 AA; 84033 MW; 1645547741D3043 CRC64;

Query Match

Best Local Similarity 9.6%; Score 87.5; DB 5; Length 759;
 Matches 28; Conservative 17; Mismatches 32; Indels 15; Gaps 5;

QY 1 SGEAGVPEKRPEDL--RCPEITTSQALPAFTTEIOASBEGADVE-----QVAPANAL 53
 Db 577 ASESTIQKRSIFDLTERCESEKNSKNSPKVNLKVSAS--PSTSEVSSIEHKFPV---- 630
 QY 54 PARSEAAVQVPIGISQVRMNSKEKKDLGTL 85
 Db 631 PARSRVAKVPIIGAKKGAALN--EVKVVHY 660

RESULT 6

Q9BL62 PRELIMINARY; PRT; 290 AA.
 AC Q9BL62;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical 33.3 kDa protein.
 GN Y65B4A.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=BRISTOL N2;
 MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Maupin R.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC024844; AAK29959.2; -.
 KW Hypothetical protein.
 SQ SEQUENCE 290 AA; 33264 MW; 1BD039FDBA67EA77 CRC64;

```

Query Match 9.4%; Score 86; DB 5; Length 290;
Best Local Similarity 18.8%; Pred. No. 4.7;
Matches 30; Conservative 29; Mismatches 55; Indels 46; Gaps 4;

QY 7 PEKPCEDLRCPETTSQALPAFTTEIQASGPGADVQVAPANALPARSEAAVQPI 66
   |||
   |||
Db 95 PEK-----KLTIDDEKETOYHSGSHSENISPNKSPFLKN-LKLTDELR 137
   |||
   |||
QY 67 GISORVEMNSKEKKDLGTGLGVLTVMVILIAIGAGIILGSYKRGDLKEQHDQKCE 126
   |||
   |||
Db 138 ENAGARPAEAKERTL-----EKDKSEMARRRREKED 172
   |||
   |||
QY 127 REMQRITLPLSAFTNPCEIVDEKTVVHTSQTVPDPOEG 166
   |||
   |||
Db 173 KETRRIEREISATNSKCELY---TFCHVGKTVIDWHG 208
   |||
   |||

RESULT 7
Q9CGO PRELIMINARY; PRT; 1165 AA.
Q9CGO
AC Q9CGO;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE KIAA1703 protein (Fragment).
DE KIAA1703.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
   |||
   |||
R1 SEQUENCE FROM N.A.
RX MEDLINE=21082932; PubMed=11214970;
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Ref. 7:347-355(2000).
DR EMBL; AB051490; BAB21794.1; -.
DR HSSP; P08047; 1SP2.
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 12.
DR SMART; SM00355; ZNF_C2H2; 12.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 7.
RW DNA-binding; Metal-binding; Zinc-finger.
FT NON TER
SQ SEQUENCE 1165 AA; 126980 MW; A37B8A9701F5133E CRC64;

Query Match 9.4%; Score 86; DB 4; Length 1165;
Best Local Similarity 24.7%; Pred. No. 26;
Matches 54; Conservative 30; Mismatches 83; Indels 52; Gaps 13

QY 2 GEAGVPEK-RP-CEDLRCPETTSQALPAFTTEIQASGPGADVQVAPANALPARSE- 58
   |||
   |||
Db 921 GAGAGEEGRPGAKDVL-----QLPG--QEVSHVAADPEAPEIQMFQADSPAAVEV 972
   |||
   |||
QY 59 -AAAVQPVIGISORVAMNSKEK-----DLGTGYVL--GITMMVLI----- 96
   |||
   |||
Db 973 LTVQVHPASAMASQEAQAQVAFKMGVGLQFAYVCDTPAAGQLKVDGTVQVVSSEGAVHM 1032
   |||
   |||
QY 97 IIAIGAGIILGSYKRGDLKE---OHQDKVCEERMQRTILPLSA--FTNPCEIV---- 147
   |||
   |||
Db 1033 VAGEGQIIMQEAQSGHMDLVESDGISQIIVTELVQAMVQSSGSGFSGTHYLTTEL 1092
   |||
   |||
QY 148 -----DEKTVVHTSQTVPDPOE---GSTPLMGQAGTP 177
   |||
   |||
Db 1093 PPGVDEPGLYSHTVLETADSCELLQAGAT-LGTTEGAP 1130
   |||
   |||

RESULT 8
Q9IV14 PRELIMINARY; PRT; 1115 AA.
Q9IV14

```

```

AC Q91V14; 2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE K-C1 cotransporter.
GN
NCBI_SLC12A5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX
NCBI_Taxid=10090;
RN
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=JUS, AND ISS:
RA Ehringer M.A., Thompson J., Conroy O., Xu Y., Yang F., Camniff J.,
RA Beeson M., Gordon L., Bennett B., Johnson T.E., Sikeia J.M.;
RT "High-Throughput Sequence Identification of Gene Coding Variants
RT within Alcohol-Related QTLs."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF332064; AAK56093.1; -.
DR EMBL; AF332063; AAK56092.1; -.
DR InterPro; IPR002293; AA/re1_primeae1.
DR InterPro; IPR004842; KCL_cotransport.
DR TIRFAMS; TIGR00930; 2a30.1.
SQ
SEQUENCE 1115 AA; 123587 MW; D506AC24D8492B4 CRC64;

Query Match 9.3%; Score 85.5; DB 11; Length 1115;
Best Local Similarity 27.1%; Pred. No. 27;
Matches 32; Conservative 15; Mismatches 44; Indels 27; Gaps 4.

Oy 79 KKDGLTGVGIGITMVTITIAIGAGIILIGYSY-----KRGKDYKEQH-DQKVCEREMQ 130
Db 872 KKDLTTPFLYHLRTAEVEVENHESDIAFYETKTLVMEQSOLKQNHILTGNRERERIQ 931
Oy 131 RT-----LPLSAFTNPTC--ELVDKTVVYVHNSQTPVDPQEGSTP 169
Db 932 SITDESRSIRRKPNPRLRLNVPBEETACDNEKPEEVEVLHIDQASAPCSSESPSP 999

RESULT 9
O63633 PRELIMINARY; PRT; 1116 AA.
AC Q63633;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Furosemide-sensitive K-C1 cotransporter.
GN KC22.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX
NCBI_Taxid=10116;
RN
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE DAWLEY; TISSUE=BRAIN;
RA MEDLINE=96279171; PubMed=8663311;
RA Payne J.A., Stevenson T.J., Donaldson L.F.;
RT "Molecular characterization of a putative K-C1 cotransporter in rat
RT brain. A neuronal-specific isoform."
RL J. Biol. Chem. 271:16245-16252(1996).
RN
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE DAWLEY; TISSUE=BRAIN;
RA MEDLINE=96279170; PubMed=86633127;
RA Gillen C.M., Brill S., Payne J.A., Forbush B. III;
RT "Molecular cloning and functional expression of the K-C1 cotransporter
RT from rabbit, rat, and human. A new member of the cation-chloride
RT cotransporter family."
RL J. Biol. Chem. 271:16237-16244(1996).
DR EMBL; U55816; AAC52635.1; -.
DR InterPro; IPR002293; AA/re1_primeae1.
DR InterPro; IPR004842; KCL_cotransport.
DR InterPro; IPR000076; KCL_transport.
DR PRINTS; PRO1081; KCLTRNSPORT.

```


DR TIGR00930; 2a30; 1.
 SQ SEQUENCE 1116 AA; 123563 MW; 882FDAOD7FDE2F6 CRC64;
 Query Match 9.3%; Score 85.5; DB 11; Length 1116;
 Best Local Similarity 27.1%; Pred. No. 27;
 Matches 32; Conservative 15; Mismatches 44; Indels 27; Gaps 4;
 QY 79 KKDGLTGYVLGITMWTIIAIGAGILIGSY-----KRGDKLRQH-DQKVCERMQ 130
 DB 872 KKDGLTGYVLGITMWTIIAIGAGILIGSY-----KRGDKLRQH-DQKVCERMQ 931
 QY 131 RIT-----LPLSAFTNPTC---EIVDEKTVVHTSOTPVDPQSGSTP 169
 DB 932 SITEBSGSIIRKKNPANTRLINVPETACDNEXKEPEEVQLHIDQSPSCPSSPSP 989
 RESULT 10
 Q8U3P9 PRELIMINARY; PRT; 607 AA.
 AC Q8U3P9;
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
 DE Hypothetical protein PF0407.
 GN PF0407.
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=2261;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the Pyrococcus furiosus genome."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE010163; AAL80531.1; -
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 607 AA; 68721 MW; 0D125AC09DE4A9AD CRC64;
 Query Match 9.2%; Score 84.5; DB 17; Length 607;
 Best Local Similarity 28.8%; Pred. No. 16;
 Matches 36; Conservative 10; Mismatches 38; Indels 41; Gaps 6;
 QY 24 ALPAFTTIOASBEGPADDEVQVAPANAL--PARSEAAVQ-----PYGISQVRMNSK 77
 DB 431 AVIPEFTREFTNTSPGHEHIVPEPPGNEIYTPASNTLVVEVTLVTEI----- 481
 QY 78 EKDDGLTGYVLGITMWTIIAIGAGILIGSYKRGKDL-----KEQHDQK-VCE 126
 DB 482 ----LGIIG-----IVITAVFLKRSRQGTGFGPIKPELLTQDHDHTKQISV 526
 QY 127 REMOR 131
 DB 527 REAYR 531
 RESULT 11
 Q93J57 PRELIMINARY; PRT; 517 AA.
 AC Q93J57;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DE Putative integral membrane protein.
 GN SC00987 OR SCBAC19F3.14.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,

RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch B., Rajandream M.A., Rutherford K., Rutter S., Taylor K.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorek A., Woodward J., Barrall B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2)";
 RL Nature 417:141-147(2002).
 DR EMBL; AL596102; CAC44337.1; -
 SQ SEQUENCE 517 AA; 54935 MW; 8C588592126A1F0 CRC64;
 Query Match 9.1%; Score 83.5; DB 16; Length 517;
 Best Local Similarity 28.7%; Pred. No. 16;
 Matches 33; Conservative 15; Mismatches 40; Indels 27; Gaps 6;
 QY 21 TSOALPAFTT--EIQASBEGPAD-ENVQVAPAN-----ALPARSEAAVQPV 65
 DB 323 TSGGLPAGLTVQVLEKARAAGLDGFEISTPADDRSTWSVAQTDGLPVRMDQAAVDPH 382
 QY 66 IG-ISCVRANMSK-----KKDGLT-GYVLGITMWTIIAIGAGILIGY 108
 DB 383 TGEVTAARVEMADHPVALKLTSLGVRAHMGTLFGLANQIVLAIVLGLTGTIVLG 437
 RESULT 12
 Q9PA31 PRELIMINARY; PRT; 733 AA.
 AC Q9PA31;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Primosomal protein N.
 GN XF2689.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Adreu F.A., Acencio M.,
 RA Alverenga R., Alves L.M.C., Araya J.E., Bata G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
 RA Colaço N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furlan L.R.,
 RA Garner M., Goldman G.H., Goldman W.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohnselt J.D., Junqueira M.L., Kemper E.L., Kitejima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nham A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A.Jr., Pasquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodriques V., de Rosa A.U.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Sanceli R.V., Sawaeski H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidants J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa."
 RL Nature 406:151-159(2000).
 CC -1- SIMILARITY: TO HELICASE C-TERMINAL DOMAIN.

DR EMBL; AE004075; AAF85486.1; -
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR005259; PriA.
 DR Pfam; PF00271; helicase_C.1.
 DR SMART; SM00490; HELIC_C.1.
 DR TIGRfam; TIGR00595; priA.1.
 KW ATP-binding; Helicase; Complete proteome.
 SQ SEQUENCE 733 AA; 81121 MW; E1341AFDDE759078 CRC64;

Query Match 9.1%; Score 83.5; DB 16; Length 733;
 Best Local Similarity 22.9%; Pred. No. 25;
 Matches 35; Conservative 17; Mismatches 52; Indels 49; Gaps 3;

QY 9 KRCEDELRCEPTTSQALPAFTTEIOEASE-----GPGADEYQVAPAN 51
 DB 181 KKPPTVSLIIQTSTNTVPVANTEQKAVETLNANNGFKTYLDGVTGSKTEVYLOAIAT 240
 QY 52 ALPARSEAAAVQPVIGISQVRNMSKEKKDL-----GTLGYV 88
 DB 241 CLAAKQALVLPVEIGLTPQLTNFPHALGI PVHALSKLADNERARVMAARCGEARLV 300
 QY 89 LGITMVIILAIAGIILGYSGKRGDKLKEQHD 121
 DB 301 LGTRSAVFTPLPHAGLIT-----VDREHD 324

RESULT 13

Q95077 PRELIMINARY; PRT; 1553 AA.

AC Q95077; 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical 173.2 kDa protein.
 GN C32E12.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=9069613; PubMed=9851916;
 RA Waterston R.;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Wilcox L.;
 RT "The sequence of C. elegans cosmid C32E12."
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U80032; AL16309.2; -
 KW Hypothetical protein.
 SQ SEQUENCE 1553 AA; 17196 MW; 02EAA041B6620CCE CRC64;

Query Match 9.1%; Score 83; DB 5; Length 1553;
 Best Local Similarity 18.1%; Pred. No. 70;
 Matches 36; Conservative 33; Mismatches 76; Indels 54; Gaps 5;

QY 3 EAGVPRKRCEDLRCE-----TTSQALPAFTTIQESBSPGADDEVQV 46
 DB 354 KSSVEETKTFETEKSEPVQKATEPEKIDNFSPTVTPRGVPLVDTSSDEBEDVEI 411
 QY 47 FAPANALPARSEAAAVQPVIGISQVRNMSKEKKDLGTLGYLGTMMVIIAIGAGIIL 106
 DB 412 IKPFIQ-PEKASEVLPVPIKVVSSIKINGIVKDV----- 446

QY 107 GYSYKRGDLKEQHDQKVCEREMQRTILPLSAFTNPCEIVDEKTVV----HTSQTPV- 161
 DB 447 -----EKTEDKKKEVDKRVATNIDVPTLPISFETSKVPEMLISSAVKVLNHDGSIIPH 502
 QY 162 -----DROEGSTPLMGQAGT 176
 DB 503 EKSDDPQEFSTLEKSESSEDT 521

RESULT 14

Q9L128 PRELIMINARY; PRT; 774 AA.

AC Q9L128; 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE ESTs D23839 (R0339).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 RT clone: P0708602."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP001539; BAA92910.1; -
 DR InterPro: IPR004087; KH dom.
 DR InterPro: IPR004088; KH_type_1.
 DR Pfam; PF00013; KH-domain; 4.
 DR SMART; SM00322; KH; 4.
 DR PROSITE; PSS0084; KH type 1; 4.
 SQ SEQUENCE 774 AA; 84066 MW; 5992CCFE0F74B856 CRC64;

Query Match 9.0%; Score 82; DB 10; Length 774;
 Best Local Similarity 23.0%; Pred. No. 37;
 Matches 29; Conservative 26; Mismatches 59; Indels 12; Gaps 4;

QY 41 ADEVQVAPANALP-----ARSEAAVQPVIGISQVRNMSKEKKDLGTLGYLGTMM 94
 DB 234 ADDSSVAKDINSREHQLBEKEMPLAVKAILVFPDRIFVNEHNGTGASGRNHSLSR 293
 QY 95 VIIAIGAGIILGYSGKRGDLKEQHDQKVCEREMQRTILPLSAF-TNPTEIVDEK-T 151
 DB 294 LTVLDSSGVGMILG---KNGSVIKQMSSTDSCEIRVSKDKLPLCALPRDELQILHOKIIT 350
 QY 152 VVVHTS 157
 DB 351 YSLHIS 356

RESULT 15

Q8W4W9 PRELIMINARY; PRT; 535 AA.

AC Q8W4W9; 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Transmembrane protein.
 GN pm3.
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4113;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. DESIREE; TISSUE=LEAF;
 RX MEDLINE=21576513; PubMed=11719809;

RA Rausch C., Daram P., Bruner S., Janas J., Laloi M., Leggewie G.,
 RA Amrhein N., Bucher M.,
 RT "A phosphate transporter expressed in arbuscule-containing cells in
 RT potato";
 RL Nature 414:462-466 (2001).
 DR EMBL; AJ318822; CAC87043.1; -.
 DR InterPro; IPR004738; Phos_permease.
 DR InterPro; IPR003662; sub_transporter.
 DR Pfam; PF00083; sugar_tr; 1.
 DR TIGRPFAM; TIGR00887; 2A0109; 1.
 SQ SEQUENCE 535 AA; 58565 MW; 94AF3CFC1839B1C1 CRC64;

Query Match 8.9%; Score 81.5; DB 10; Length 535;
 Best Local Similarity 25.8%; Pred. NO. 26;

Matches 32; Conservative 15; Mismatches 50; Indels 27; Gaps 5;

QY 35 ASBGPQADAEVQVFPANALPAR-----SEAAVQPVIGISQRYMNSKEKKDL 82
 DB 417 ANFGPNATTFV--PAEIFPARLRSTCHGISAAAGAGAIYGAFGLYAAQSTDPKVD- 473
 QY 83 GTIGYVIGITMTMTITTAIGAGIILGYSY-----KRGKDLKEQHDQVCEREMQRT---- 132
 DB 474 --AGYPTGIGKVALIVLCVNFGLMFTLLVPSKSGKSLSEMSKENGEEMTKVENAQ 531

QY 133 TLPL 136
 DB 532 TTPV 535

Search completed: April 7, 2003, 09:21:41
 Job time : 53.0062 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:24 ; Search time 33.4927 Seconds

(without alignments)
712.151 Million cell updates/sec

Title: US-10-057-951-2_COPY_85_263

Perfect score: 916
Sequence: 1 SGEAGVPRKPCEDRCPEPT.....PVDQEGSTPLMGAGTPCA 179

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 916 | 100.0 | 263 | 20 | AAV05219 |
| 2 | 916 | 100.0 | 263 | 22 | AAW87769 |
| 3 | 916 | 100.0 | 263 | 22 | AAE00300 |
| 4 | 913 | 99.7 | 263 | 21 | AAW43237 |
| 5 | 913 | 99.7 | 263 | 23 | AAU086149 |
| 6 | 902 | 98.5 | 263 | 22 | AAW93748 |
| 7 | 704.5 | 76.9 | 286 | 20 | AAV05220 |
| 8 | 322 | 35.2 | 66 | 22 | ABB37905 |
| 9 | 322 | 35.2 | 66 | 22 | ABB23159 |
| 10 | 322 | 35.2 | 66 | 22 | AAW58537 |

| | | | | | | |
|----|------|------|------|----|----------|---------------------|
| 11 | 322 | 35.2 | 66 | 22 | AAW1037 | Human bone marrow |
| 12 | 322 | 35.2 | 66 | 22 | AAW18800 | Peptide #5234 enco |
| 13 | 322 | 35.2 | 66 | 22 | AAW13314 | Peptide #5351 enco |
| 14 | 322 | 35.2 | 66 | 22 | ABG40828 | Human peptide enco |
| 15 | 95 | 10.4 | 163 | 22 | AAU58844 | Propionibacterium |
| 16 | 94 | 10.3 | 8805 | 22 | ABW67112 | Drosophila melanog |
| 17 | 86.5 | 9.4 | 507 | 20 | AAW34756 | C. pneumoniae prot |
| 18 | 80.5 | 8.8 | 769 | 21 | AAW53896 | A. Neisseria mening |
| 19 | 80 | 8.7 | 1338 | 21 | ABW07823 | Human notch agonis |
| 20 | 79.5 | 8.7 | 769 | 21 | AAW53897 | A. Neisseria mening |
| 21 | 79 | 8.6 | 1212 | 19 | AAW44299 | Human serrate 2. |
| 22 | 78.5 | 8.6 | 552 | 22 | ABW1591 | Drosophila melanog |
| 23 | 77.5 | 8.5 | 452 | 22 | AAW82370 | S. epidermidis ope |
| 24 | 77.5 | 8.5 | 463 | 23 | ABW3128 | Staphylococcus epi |
| 25 | 77.5 | 8.5 | 595 | 23 | AAW78088 | Human CD30 protein |
| 26 | 77 | 8.4 | 769 | 21 | AAW53895 | A. Neisseria mening |
| 27 | 76.5 | 8.4 | 425 | 21 | AAW92370 | p75-NTR (neurotrop |
| 28 | 76.5 | 8.4 | 425 | 23 | AAE21671 | Rat neurotrophic r |
| 29 | 76.5 | 8.4 | 461 | 19 | AAW75391 | Human liver glycop |
| 30 | 76.5 | 8.4 | 501 | 19 | AAW75391 | Human liver glycop |
| 31 | 76.5 | 8.4 | 556 | 22 | ABG00252 | Novel human diagno |
| 32 | 76.5 | 8.4 | 556 | 22 | ABG00603 | Novel human diagno |
| 33 | 76.5 | 8.4 | 556 | 22 | ABG03054 | Novel human diagno |
| 34 | 76.5 | 8.4 | 1155 | 22 | AAW40061 | Human polypeptide |
| 35 | 75.5 | 8.2 | 219 | 23 | ABW40806 | Staphylococcus epi |
| 36 | 75.5 | 8.2 | 219 | 23 | ABW31724 | Peptide #4375 enco |
| 37 | 75.5 | 8.2 | 276 | 22 | ABW36945 | Peptide #4451 enco |
| 38 | 75.5 | 8.2 | 276 | 22 | AAW70088 | Human bone marrow |
| 39 | 75.5 | 8.2 | 276 | 22 | AAW17915 | Peptide #4349 enco |
| 40 | 75.5 | 8.2 | 276 | 22 | AAW30426 | Peptide #4463 enco |
| 41 | 75.5 | 8.2 | 276 | 22 | AAW05567 | Peptide #4249 enco |
| 42 | 75.5 | 8.2 | 276 | 23 | ABG39722 | Human peptide enco |
| 43 | 75.5 | 8.2 | 665 | 22 | ABG1591 | Novel human diagno |
| 44 | 75 | 8.2 | 595 | 14 | AAW35478 | Lymphocyte activat |
| 45 | 75 | 8.2 | 595 | 18 | AAW24017 | Human soluble CD30 |

ALIGNMENTS

| | |
|----------|--|
| RESULT 1 | |
| AAV05219 | |
| ID | AAV05219 standard; protein; 263 AA. |
| XX | |
| AC | AAV05219; |
| XX | |
| DT | 17-JUN-1999 (first entry) |
| XX | |
| DE | Kringel protein sequence. |
| XX | |
| KW | Kringel; cancer; inflammation; autoimmunity; rheumatoid arthritis; AIDS; |
| KW | CNS inflammation; cerebellar degeneration; Alzheimer's disease; asthma; |
| KW | Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis; |
| KW | neurological abnormality; ischemia reperfusion injury; ischemic injury; |
| KW | cardiovascular disease; kidney disease; liver disease; aplastic anemia; |
| KW | myocardial infarction; hypotension; hypertension; allergy; infection; |
| KW | myelodysplastic syndrome; hematologic abnormality; diagnosis; therapy; |
| KW | male pattern baldness. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | W09911788-A1. |
| XX | |
| PD | 11-MAR-1999. |
| XX | |
| PR | 02-SEP-1998; 98WO-US18270. |
| XX | |
| PR | 01-SEP-1998; 98US-0144889. |
| PR | 02-SEP-1997; 97US-0056032. |
| XX | |
| PA | (SMIK) SMITHKLINE BEECHAM CORP. |
| XX | |
| PI | Albone EF, Kikly KK; |

```
XX WPI: 1999-214707/18.
DR N-PSDB; AAX28354.
XX
XX New kringlet polypeptides and polynucleotides
XX
XX Claim 1; Page 31-32; 42pp; English.
XX
XX This sequence is a kringlet polypeptide of the invention.
XX The kringlet polypeptides (I) are used to screen for agonists and
XX antagonists. Agonists are used to treat subjects in need of enhanced
XX activity or expression of (I). Antagonists are used to treat subjects
XX having need to inhibit the activity or expression of (I). The methods can
XX be used to treat conditions such as cancer, inflammation, autoimmunity,
XX allergy, asthma, rheumatoid arthritis, CNS inflammation, cerebellar
XX degeneration, Alzheimer's disease, Parkinson's disease, multiple
XX sclerosis, amyotrophic lateral sclerosis, head injury damage and other
XX neurological abnormalities, ischaemia reperfusion injury, cardiovascular
XX disease, kidney disease, liver disease, ischaemic injury, myocardial
XX infarction, hypertension, AIDS, myelodysplastic syndromes
XX and other haematologic abnormalities, aplastic anaemia, male pattern
XX baldness, and bacterial, fungal, protozoan and viral infections. The
XX kringlet polypeptides may also be used to generate antibodies.
XX Determining the presence or absence of mutations in, and analysing for
XX the presence or absence of expression of, kringlet polynucleotides can be
XX used to diagnose a disease or susceptibility to a disease related to
XX expression or activity of kringlet proteins. The polynucleotides may also
XX be used for chromosome identification, and mapping.
XX
XX Sequence 263 AA;
SQ
XX
XX Query Match 100.0%; Score 916; DB 20; Length 263;
XX Best Local Similarity 100.0%; Pred. No. 2,3e-89;
XX Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 SGEAGVPEKRPCEDLRCPEETTSQALPAFTTEIOEASSEGGADEVQVAPANALPARSEAA 60
XX DB 85 SGEAGVPEKRPCEDLRCPEETTSQALPAFTTEIOEASSEGGADEVQVAPANALPARSEAA 144
XX
XX 61 AVQPIGIGSORVRNMSKKKDLGTYLGTMMVITIIAIGAGIILGYSTYKRGDKLEQH 120
XX DB 145 AVQPIGIGSORVRNMSKKKDLGTYLGTMMVITIIAIGAGIILGYSTYKRGDKLEQH 204
XX
XX 121 DQKVCEREMQRTITPLSAFTNPCTCEIVDEKTVVHTSQTVPDPOEGSTPIPMGAGTPGA 179
XX DB 205 DQKVCEREMQRTITPLSAFTNPCTCEIVDEKTVVHTSQTVPDPOEGSTPIPMGAGTPGA 263
XX
XX RESULT 2
XX AAW87769
XX ID AAW87769 standard; Protein; 263 AA.
XX
XX AC AAW87769;
XX
XX 29-MAR-1999 (first entry)
XX
XX Human tissue plasminogen activator-like protease t-PALP.
XX
XX Tissue plasminogen activator-like protease; t-PALP; human;
XX circulatory system-related disorder; blood clotting; stroke;
XX thrombosis; peripheral arterial occlusion; pulmonary embolism;
XX myocardiothrombosis; diagnosis; therapy.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..21
XX FT /label= sig_peptide
XX FT 22..263
XX FT /label= Mat_protein
XX FT Domain 25..84
XX FT /note= "kringle domain"
XX FT Domain 85..263
```

```
FT Peptide /note= "protease domain"
FT 22..31
FT /note= "epitope-bearing region"
FT 35..44
FT /note= "epitope-bearing region"
FT 71..81
FT /note= "epitope-bearing region"
FT 91..107
FT /note= "epitope-bearing region"
FT 119..128
FT /note= "epitope-bearing region"
FT 138..147
FT /note= "epitope-bearing region"
FT 155..167
FT /note= "epitope-bearing region"
FT 193..203
FT /note= "epitope-bearing region"
FT 206..215
FT /note= "epitope-bearing region"
FT 227..237
FT /note= "epitope-bearing region"
FT 243..252
FT /note= "epitope-bearing region"
XX
XX W09854199-A1.
XX
XX 03-DEC-1998.
XX
XX 27-MAY-1998; 98WO-US10728.
XX
XX 28-MAY-1997; 97US-0048000.
XX
XX (HUMA-) HUMAN GENOMEG SCI INC.
XX
XX Ebner R, Moore PA, Ruben SM;
XX
XX WPI: 1999-070207/06.
XX DB N-PSDB; AAV99636.
XX
XX New tissue plasminogen activator-like protease - useful in the
XX diagnosis and treatment of circulatory system-related disorders
XX
XX Claim 1; Page 56-57; 76pp; English.
XX
XX This is the amino acid sequence of tissue plasminogen activator-like
XX protease (t-PALP), a novel member of the serine protease family
XX that shares sequence homology to human tissue plasminogen activator
XX (see AAW87770). The t-PALP sequence was deduced from a cDNA clone
XX (see AAV99636) derived from activated monocytes. The 2.5 kb t-PALP
XX message has also been detected in heart, brain, lung, placenta,
XX liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate,
XX testis, ovary, small intestine, colon and peripheral blood
XX leukocytes. Isolated nucleic acids encoding amino acids -21 to
XX 242, -20 to 242, 1-242, 4-63 (kringle domain) and 64-242 (protease
XX domain) of t-PALP, or encoding epitope-bearing portions of t-PALP,
XX are also claimed, as are recombinant vectors, host cells, and
XX methods for producing t-PALP polypeptides. t-PALP may be used to
XX detect and treat disorders related to the circulatory system, and
XX to identify agonists and antagonists of t-PALP activity. The
XX homology between t-PALP and tPA indicates that t-PALP may be
XX involved in the regulation of normal and abnormal clotting
XX in e.g. stroke, deep-vein thrombosis, peripheral arterial
XX occlusion, pulmonary embolism and myocardiothrombosis.
XX
XX Sequence 263 AA;
SQ
XX
XX Query Match 100.0%; Score 916; DB 20; Length 263;
XX Best Local Similarity 100.0%; Pred. No. 2,3e-89;
XX Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 SGEAGVPEKRPCEDLRCPEETTSQALPAFTTEIOEASSEGGADEVQVAPANALPARSEAA 60
XX DB 85 SGEAGVPEKRPCEDLRCPEETTSQALPAFTTEIOEASSEGGADEVQVAPANALPARSEAA 144
```



```

PR 05-JAN-2000; 2000MO-US00219.
XX
XX (GETH ) GENENTECH INC.
XX
XX Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
PI Martens SA, Pan J, Pitti RM, Roy MA, Smith V, Stone DM;
PI Watanabe CK, Wood WI;
XX
XX WPI; 2002-205567/26.
DR N-PEDB; ABR40275.
XX
XX Thirty five nucleic acids encoding PRO polypeptides, useful for
PT treating benign or malignant tumours, leukaemias and lymphoid
PT malignancies, inflammatory, angiogenic and immunologic disorders -
XX
XX Claim 61; Fig 44; 302pp; English.
XX
XX The present invention relates to the isolation of novel human PRO
CC polypeptides and the polynucleotide sequences encoding them. The
CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are
CC useful for treating benign or malignant tumors (e.g. renal, kidney,
CC bladder, breast, etc), leukaemias and lymphoid malignancies, other
CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,
CC macrophagal, stromal and blastoclastic disorders, inflammatory, immune
CC and angiogenic disorders. The polynucleotide sequences are also
CC useful in gene therapy. AAU86128-AAU86162 represent the human PRO
CC polypeptides of the invention.
XX
XX Sequence 263 AA:
SQ
XX
XX Query Match 99.7%; Score 913; DB 23; Length 263;
XX Best Local Similarity 99.4%; Pred. No. 4,8e-89;
XX Matches 178; Conservative 1; Mismatches 0; Indels 0; Gaps
XX 0;
XX
XX 1 SGEAGVPERPCEDLRCPETTSQALPAFTTETIOEASEGPGADVEQVAPANMLPARSEAA 60
DB 85 SGEAGVPERPCEDLRCPETTSQALPAFTTETIOEASEGPGADVEQVAPANMLPARSEAA 144
XX
XX 61 AAGPVIGISQRANMSKEKKDLGTLGYVIGITMMVVIILAGAGIILGYGKRGKDLKEQH 120
DB 145 AAGPVIGISQRANMSKEKKDLGTLGYVIGITMMVVIILAGAGIILGYGKRGKDLKEQH 204
XX
XX 121 DQKVCEREMQRTLLPLSAFTNPTCEIVDEKTVVWHTSQTPVDPOEGSTRPLMGQAGTPGA 179
DB 205 DQKVCEREMQRTLLPLSAFTNPTCEIVDEKTVVWHTSQTPVDPOEGSTRPLMGQAGTPGA 263
XX
XX RESULT 6
XX AAM93748
XX ID AAM93748 standard; Protein; 263 AA.
XX
XX AAM93748;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polypeptide, SEQ ID NO: 3727.
XX
XX Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
XX Homo sapiens.
XX
XX EP1130094-A2.
XX
XX 05-SEP-2001.
XX
XX 07-JUL-2000; 2000EP-0114089.
XX
XX 08-JUL-1999; 99JP-0194486.
XX 11-JAN-2000; 2000JP-0118774.
XX 02-MAY-2000; 2000JP-0183765.
XX
XX (HELI-) HELIX RES INST.
XX

```

| | | |
|---------------------------|---|--|
| PI | Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y, | |
| PI | Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H, | |
| DR | WPI, 2001-524255/58. | |
| XX | | |
| DR | N-PSDB, AAK94700. | |
| XX | | |
| PT | 830 Primers useful for synthesizing full length cDNA clones and their | |
| PT | use in genetic manipulation - | |
| PS | Claim 8, SEQ ID NO 3727, 1380bp + sequence listing; English. | |
| XX | | |
| CC | The invention relates to primers for synthesizing full length cDNA | |
| CC | clones. 830 cDNA molecules encoding a human protein have been | |
| CC | isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA | |
| CC | molecules have been determined. Primers for synthesizing the full length | |
| CC | cDNA are useful for clarifying the function of the protein encoded by | |
| CC | the cDNA. The full length clones were obtained by construction of full | |
| CC | length enriched cDNA libraries that were synthesised by the oligo-capping | |
| CC | method. The primers enable the production of the full length cDNA easily | |
| CC | without any special methods. The present sequence is a polypeptide | |
| CC | encoded by a full length human cDNA of the invention. | |
| CC | Note: The sequence data for this patent did not form part of the printed | |
| CC | specification, but was obtained in CD-ROM format directly from EPO. | |
| XX | | |
| XX | Sequence 263 AA; | |
| Query Match | 98.5%; Score 902; DB 22; Length 263; | |
| Best Local Similarity | 98.9%; Pred. No. 7.2e-89; | |
| Matches 177; Conservative | 0; Mismatches 2; Indels 0; Gaps 0; | |
| QY | 1 SGEAGVPEKRPCEDLRCPTTSQALPAFTTEIQEASBGPADDEVQFAPANALPARSEAA 60 | |
| DB | 85 SGEAGVPEKRPCEGRCPETTSQALPAFTTEIQEASBGPADDEVQFAPANALPARSEAA 144 | |
| QY | 61 AVQPIGTISQRRNMSKSKKDLGTGYVLGITMVTITIALGAGILIGYSYKRGKDLKEQH 120 | |
| DB | 145 AVQPIGTISQRRNMSKSKKDLGTGYVLGITMVTITIALGAGILIGYSYKRGKDLKEQH 204 | |
| QY | 121 DQKVEREMORTTLPISAFNTPTCETIVDEKTVVHTSOTPVDPQBSGTPLMGQAGTPGA 179 | |
| DB | 205 DQKVEREMORTTLPISAFNTPTCETIVDEKTVVHTSOTPVDPQBSGTPLMGQAGTPGA 263 | |
| RESULT 7 | | |
| AAV05220 | | |
| ID | AAV05220 standard; Protein; 286 AA. | |
| XX | AAV05220; | |
| AC | | |
| XX | | |
| DT | 17-JUN-1999 (first entry) | |
| XX | | |
| DE | Kringel1 protein sequence. | |
| XX | | |
| KW | Kringel1; cancer; inflammation; autoimmunity; rheumatoid arthritis; AIDS; | |
| KW | CNS inflammation; cerebellar degeneration; Alzheimer's disease; asthma; | |
| KW | Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis; | |
| KW | neurological abnormality; ischaemia reperfusion injury; ischaemic injury; | |
| KW | cardiovascular disease; kidney disease; liver disease; aplastic anaemia; | |
| KW | myocardial infarction; hypotension; hypertension; allergy; infection; | |
| KW | myelodysplastic syndrome; haematologic abnormality; diagnosis; therapy; | |
| KW | male pattern baldness. | |
| XX | | |
| OS | Homo sapiens. | |
| XX | | |
| PN | W09911788-A1. | |
| XX | | |
| PD | 11-MAR-1999. | |
| XX | | |
| PF | 02-SEP-1998; 98WO-US18270. | |
| XX | | |
| XX | 01-SEP-1998; 98US-0144889. | |
| PR | 02-SEP-1997; 97US-0056032. | |

PA (SMIK) SMITHKLINE BEECHAM CORP.
XX Albione EF, Kikily KK;
XX WPI; 1999-214707/18.
DR N-PSDB; AAX28355.
XX
PT New kringlet polypeptides and polynucleotides
XX
PS Claim 14; Page 33; 42pp; English.
XX
CC This sequence is a kringlet polypeptide of the invention.
CC The kringlet polypeptides (I) are used to screen for agonists and
CC antagonists. Agonists are used to treat subjects in need of enhanced
CC activity or expression of (I). Antagonists are used to treat subjects
CC having need to inhibit the activity or expression of (I). The methods can
CC be used to treat conditions such as cancer, inflammation, cerebellar
CC degeneration, Alzheimer's disease, Parkinson's disease, multiple
CC sclerosis, amyotrophic lateral sclerosis, head injury damage and other
CC neurological abnormalities, ischaemia reperfusion injury, cardiovascular
CC disease, kidney disease, liver disease, ischaemic injury, myocardial
CC infarction, hypertension, AIDS, myelodysplastic syndromes
CC and other haematologic abnormalities, aplastic anaemia, male pattern
CC baldness, and bacterial, fungal, protozoan and viral infections. The
CC kringlet polypeptides may also be used to generate antibodies.
CC Determining the presence or absence of mutations in, and analysing for
CC the presence or absence of expression of, kringlet polynucleotides can be
CC used to diagnose a disease or susceptibility to a disease related to
CC expression or activity of kringlet proteins. The polynucleotides may also
CC be used for chromosome identification, and mapping.
XX
SQ Sequence 286 AA;

Query Match 76.9%; Score 704.5; DB 20; Length 286;
Best Local Similarity 82.4%; Pred. No. 9.9e-67;
Matches 145; Conservative 5; Mismatches 25; Indels 1; Gaps 1;

QY 1 SGEAGVPEKRPCEBDRPETSQALPAF-TTEIOEASGEGADEVOVFAPANALPARSEA 59
DB 85 SGEAGVPEKRPCEBDRPETSQALPAFHDRNSKRLEGGADEVQVFAPANALPARSEA 144
QY 60 AAVQPVIGISQRYVMSKEKKDLGTLGYVLITMVMVITIIAGIILGYSYKRGKDLKEQ 119
DB 145 ALLOPVIGISQRYVDEQGEKRGNSGLRAGHHDSIILIGAGIILGYSYKRGKDLKEQ 204
QY 120 HDQKVCEREMQRTILPLFAFTNPCEIVDEKTVVHTSQTPVDPOEGSTPLMGOAG 175
DB 205 HDQKVCEREMQRTILPLFAFTNPCEIVDEKTVVHTSQTPVDPOEGSTPLMGOAG 260

RESULT 8
ABB37905
ID ABB37905 standard; Peptide; 66 AA.
XX
AC ABB37905;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #5411 encoded by human foetal liver single exon probe.
XX
KM Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001MO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-02344687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human fetal liver -
XX
PS Claim 27; SEQ ID NO 30540; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 66 AA;

Query Match 35.2%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 8.9e-27;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 TTSQLPAPFTTEIOBASGEGADEVOVFAPANALPARSEAAAVPVIGISQRYVMSKEK 79
DB 1 TTSQLPAPFTTEIOBASGEGADEVOVFAPANALPARSEAAAVPVIGISQRYVMSKEK 60
QY 80 KDLGTL 85
DB 61 KDLGTL 66

RESULT 9
ABB23159
ID ABB23159 standard; Protein; 66 AA.
XX
AC ABB23159;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #5158 encoded by probe for measuring heart cell gene expression.
XX
KM Human; gene expression; heart; microarray; vascular system;
KM cardiovascular disease; hypertension; cardiac arrhythmia;
KM congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001MO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-02344687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

```
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PT WPI; 2001-48899/53.
XX DR
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX PS Claim 15; SEQ ID No 24929; 530pp; English.
XX CC The present invention relates to single exon nucleic acid probes for
XX CC measuring human gene expression in a sample derived from human heart (see
XX CC ABA21533-ABA41305). The present sequence is a protein encoded by one such
XX CC probe. The probes may be used for predicting, measuring and displaying
XX CC gene expression in samples derived from the human heart via microarrays.
XX CC By measuring gene expression, the probes are useful for predicting,
XX CC diagnosing, grading, staging, monitoring and prognosticating diseases of the
XX CC human heart and vascular system e.g. cardiovascular disease,
XX CC hypertension, cardiac arrhythmias and congenital heart disease.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 66 AA;
XX
XX Query Match 35.2%; Score 322; DB 22; Length 66;
XX Best Local Similarity 100.0%; Pred. No. 8.9e-27;
XX Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 20 TTSQLPAPFTTEIOEASGPGADEVOYFAPANALPARSEAAAVQPVIGISORVRMSKKEK 79
DB 1 TTSQLPAPFTTEIOEASGPGADEVOYFAPANALPARSEAAAVQPVIGISORVRMSKKEK 60
QY 80 KDLGTL 85
DB 61 KDLGTL 66
DB
XX
XX RESULT 10
XX AAM58537
XX ID AAM58537 standard; Protein; 66 AA.
XX AC
XX AAAM58537;
XX DT 05-NOV-2001 (first entry)
XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30642.
XX XX
XX KM Human; brain expressed exon; gene expression analysis; probe;
XX KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX KM epilepsy; cancer.
XX OS Homo sapiens.
XX XX
XX PN WO200157275-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00667.
XX XX
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PS (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PT WPI; 2001-483446/52.
```

```
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX PT brains -
XX PS Example 4; SEQ ID NO: 30642; 650pp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC brain. They can be used to measure gene expression in brain cell samples,
XX CC which may enable the diagnosis and improved treatment of nervous system
XX CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX CC epilepsy and cancers. The present sequence is a protein encoded by one of
XX CC the probes of the invention.
XX SQ Sequence 66 AA;
XX
XX Query Match 35.2%; Score 322; DB 22; Length 66;
XX Best Local Similarity 100.0%; Pred. No. 8.9e-27;
XX Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 20 TTSQLPAPFTTEIOEASGPGADEVOYFAPANALPARSEAAAVQPVIGISORVRMSKKEK 79
DB 1 TTSQLPAPFTTEIOEASGPGADEVOYFAPANALPARSEAAAVQPVIGISORVRMSKKEK 60
QY 80 KDLGTL 85
DB 61 KDLGTL 66
DB
XX
XX RESULT 11
XX AAM71037
XX ID AAM71037 standard; Protein; 66 AA.
XX AC
XX AAAM71037;
XX DT 06-NOV-2001 (first entry)
XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31343.
XX XX
XX KM Human; bone marrow expressed exon; gene expression analysis; probe;
XX KM microarray; cancer; leukaemia; lymphoma; myeloma.
XX OS Homo sapiens.
XX XX
XX PN WO200157276-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00668.
XX XX
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PS (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PT WPI; 2001-488900/53.
XX DR
XX PT Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human bone marrow -
XX XX
XX PS Example 4; SEQ ID NO: 31343; 658pp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC bone marrow. They can be used to measure gene expression in bone marrow
XX CC samples, which may enable the improved diagnosis and treatment of cancers
```


RESULT 14
 ABG40828
 ID ABG40828 standard; Peptide; 66 AA.
 AC ABG40828;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 30493.
 XX
 KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW primary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00665.
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024253.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX
 PS Claim 27; SEQ ID No 30493; 634pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 1614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe.
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene

CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 66 AA;
 XX
 QY Query Match 35.2%; Score 322; DB 23; Length 66;
 Db Best Local Similarity 100.0%; Pred. No. 8.9e-27;
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 20 TTSGALPATTETIQEASGSGADDEVQVFPANALPAREBAAVQVIGISGRVKNRSKEX 79
 Db 1 TTSGALPATTETIQEASGSGADDEVQVFPANALPAREBAAVQVIGISGRVKNRSKEX 60
 QY 80 KDLGTL 85
 Db 61 KDLGTL 66
 XX
 RESULT 15
 AAU55844
 ID AAU55844 standard; Protein; 163 AA.
 XX
 AC AAU55844;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #16740.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhattacha A;
 PI L'maiesmauve J, Zhang Y, Jen S, Carter D,
 XX
 DR WPI; 2001-616774/71.
 DR N-PSDB; AAS39572.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID No 17039; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:19:28 ; Search time 12.6528 Seconds

(without alignments)
864.895 Million cell updates/sec

Title: US-10-057-951-2_COPY_85_263

Perfect score: 916 1 SGEAGVPEKRPCEDLRCPET.....PVDPQESTPLMGQAGTPGA 179

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications Aa:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PC7_NEW_PUB pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/FC105_PUBCOMB pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Length | DB ID | Description |
|------------|-------|--------------|----------------------------|--------------------|
| 1 | 916 | 100.0 | 263 9 US-10-102-704-2 | Sequence 2, Appl1 |
| 2 | 916 | 100.0 | 263 9 US-10-057-951-2 | Sequence 2, Appl1 |
| 3 | 916 | 100.0 | 263 10 US-09-084-491A-2 | Sequence 2, Appl1 |
| 4 | 322 | 35.2 | 66 10 US-09-864-761-38457 | Sequence 38457, A |
| 5 | 80 | 8.7 | 1212 9 US-10-219-248-3 | Sequence 3, Appl1 |
| 6 | 80 | 8.7 | 1212 9 US-10-219-247-3 | Sequence 3, Appl1 |
| 7 | 80 | 8.7 | 1212 10 US-09-855-722-3 | Sequence 3, Appl1 |
| 8 | 80 | 8.7 | 1238 9 US-10-219-248-5 | Sequence 5, Appl1 |
| 9 | 80 | 8.7 | 1238 9 US-10-219-247-5 | Sequence 5, Appl1 |
| 10 | 80 | 8.7 | 1238 10 US-09-855-722-5 | Sequence 5, Appl1 |
| 11 | 80 | 8.7 | 1238 10 US-09-944-849-4 | Sequence 6, Appl1 |
| 12 | 77.5 | 8.5 | 1709 9 US-09-921-667-6 | Sequence 6, Appl1 |
| 13 | 77.5 | 8.5 | 1709 9 US-09-870-759-51 | Sequence 5, Appl1 |
| 14 | 76.5 | 8.4 | 425 10 US-09-748-537-14 | Sequence 14, Appl1 |
| 15 | 76.5 | 8.4 | 425 10 US-09-821-831-2 | Sequence 2, Appl1 |
| 16 | 75.5 | 8.2 | 276 10 US-09-864-761-46690 | Sequence 46690, A |
| 17 | 75 | 8.2 | 595 10 US-09-826-212-9 | Sequence 9, Appl1 |
| 18 | 75 | 8.2 | 595 10 US-09-935-727-11 | Sequence 11, Appl1 |
| 19 | 74.5 | 8.1 | 403 10 US-09-996-194-6 | Sequence 6, Appl1 |

| | | | | |
|----|------|-----|---------------------------|--------------------|
| 20 | 74.5 | 8.1 | 473 10 US-09-764-864-802 | Sequence 802, App |
| 21 | 74.5 | 8.1 | 617 9 US-09-860-670-82 | Sequence 82, Appl1 |
| 22 | 74.5 | 8.1 | 617 10 US-09-764-864-1263 | Sequence 1263, Ap |
| 23 | 73.5 | 8.0 | 1179 10 US-09-821-883-29 | Sequence 29, Appl1 |
| 24 | 73 | 8.0 | 348 10 US-09-731-872-247 | Sequence 247, App |
| 25 | 73 | 8.0 | 451 10 US-09-935-390A-26 | Sequence 26, Appl1 |
| 26 | 73 | 8.0 | 451 9 US-09-992-558-301 | Sequence 301, App |
| 27 | 73 | 8.0 | 461 9 US-09-989-293A-301 | Sequence 301, App |
| 28 | 73 | 8.0 | 461 9 US-09-989-735-301 | Sequence 301, App |
| 29 | 73 | 8.0 | 461 9 US-09-990-444-301 | Sequence 301, App |
| 30 | 73 | 8.0 | 461 9 US-09-989-730-301 | Sequence 301, App |
| 31 | 73 | 8.0 | 461 9 US-09-990-436-301 | Sequence 301, App |
| 32 | 73 | 8.0 | 461 9 US-09-991-181-301 | Sequence 301, App |
| 33 | 73 | 8.0 | 461 9 US-09-993-687-301 | Sequence 301, App |
| 34 | 73 | 8.0 | 461 9 US-09-989-734-301 | Sequence 301, App |
| 35 | 73 | 8.0 | 461 9 US-10-028-072-454 | Sequence 454, App |
| 36 | 73 | 8.0 | 461 9 US-09-997-653-301 | Sequence 301, App |
| 37 | 73 | 8.0 | 461 9 US-09-993-667-301 | Sequence 301, App |
| 38 | 73 | 8.0 | 461 9 US-10-121-049-454 | Sequence 454, App |
| 39 | 73 | 8.0 | 461 9 US-10-123-904-454 | Sequence 454, App |
| 40 | 73 | 8.0 | 461 9 US-10-140-470-454 | Sequence 454, App |
| 41 | 73 | 8.0 | 461 9 US-09-990-438-301 | Sequence 301, App |
| 42 | 73 | 8.0 | 461 9 US-09-990-562-301 | Sequence 301, App |
| 43 | 73 | 8.0 | 461 9 US-09-997-428-301 | Sequence 301, App |
| 44 | 73 | 8.0 | 461 9 US-09-997-666-301 | Sequence 301, App |
| 45 | 73 | 8.0 | 461 9 US-10-175-746-454 | Sequence 454, App |

ALIGNMENTS

RESULT 1
US-10-102-704-2
Sequence 2, Application US/10102704
Parent NO. US20020164768A1
GENERAL INFORMATION:
APPLICANT: Moore et al.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protein
FILE REFERENCE: PF378C1
CURRENT APPLICATION NUMBER: US/10/102,704
CURRENT FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/084,491
PRIOR FILING DATE: 1998-05-27
PRIOR APPLICATION NUMBER: 60/048,000
PRIOR FILING DATE: 1997-05-28
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-10-102-704-2

Query Match 100.0%; Score 916; DB 9; Length 263;
Best Local Similarity 100.0%; Pred. No. 4.9e-87;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASEGGADEVQVFAPANALPARSRA 60
DB 85 SGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASEGGADEVQVFAPANALPARSRA 144

QY 61 AVQPIYIGISQVRNNSKKKDLGTLGVTWVIAIAGGIIIGSYRKDKLKEQH 120
DB 145 AVQPIYIGISQVRNNSKKKDLGTLGVTWVIAIAGGIIIGSYRKDKLKEQH 204

QY 121 DQKCEMRMRTLPISAFNPTCEIVDEKTVVHVSQFVDPQSGSTPLMGQAGTPGA 179
DB 205 DQKCEMRMRTLPISAFNPTCEIVDEKTVVHVSQFVDPQSGSTPLMGQAGTPGA 263

RESULT 2
US-10-057-951-2
Sequence 2, Application US/10057951

Parent No. US20020177213A1
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: PF378P1
; CURRENT APPLICATION NUMBER: US/10/057,951
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 09/411,977
; PRIOR FILING DATE: 1999-10-04
; PRIOR APPLICATION NUMBER: US 09/084,491
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: US 60/048,000
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-951-2

Query Match 100.0%; Score 916; DB 9; Length 263;
Best Local Similarity 100.0%; Pred. No. 4,9e-87;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIQEASGGADEVGFAPANALPARSEAA 60
DB 85 SGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIQEASGGADEVGFAPANALPARSEAA 144
QY 61 AVOPVIGISQVRNRSKEDLGLGVIGITMVIILIGAGIILGYSYRGDKLEQH 120
DB 145 AVOPVIGISQVRNRSKEDLGLGVIGITMVIILIGAGIILGYSYRGDKLEQH 204
QY 121 DQVCEREMQRTPLPSAFNPTCEIVDEKTVVHTSQTPVDPQGSTPLMGQGTPGA 179
DB 205 DQVCEREMQRTPLPSAFNPTCEIVDEKTVVHTSQTPVDPQGSTPLMGQGTPGA 263

RESULT 3

US-09-084-491A-2
; Sequence 2, Application US/09084491A
; Patent No. US2002061576A1
; GENERAL INFORMATION:
; APPLICANT: MOORE, PAUL A.
; APPLICANT: RUBEN, STEVEN M.
; APPLICANT: EBNER, REINHARD
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/084,491A
; FILING DATE: 27-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,573
; REFERENCE/DOCKET NUMBER: PF378
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-084-491A-2

Query Match 100.0%; Score 916; DB 10; Length 263;
Best Local Similarity 100.0%; Pred. No. 4,9e-87;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIQEASGGADEVGFAPANALPARSEAA 60
DB 85 SGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIQEASGGADEVGFAPANALPARSEAA 144
QY 61 AVOPVIGISQVRNRSKEDLGLGVIGITMVIILIGAGIILGYSYRGDKLEQH 120
DB 145 AVOPVIGISQVRNRSKEDLGLGVIGITMVIILIGAGIILGYSYRGDKLEQH 204
QY 121 DQVCEREMQRTPLPSAFNPTCEIVDEKTVVHTSQTPVDPQGSTPLMGQGTPGA 179
DB 205 DQVCEREMQRTPLPSAFNPTCEIVDEKTVVHTSQTPVDPQGSTPLMGQGTPGA 263

RESULT 4

US-09-864-761-38457
; Sequence 38457, Application US/09864761
; Patent No. US2002048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David K.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29

US-09-855-722-3

Query Match 8.7%; Score 80; DB 10; Length 1212;
Best Local Similarity 22.4%; Pred. No. 13;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;

QY 2 GEAGV--PEKRPC-----EDIRC-----PETTS-----QALPAFTTEIQ 33
DB 921 GEGCAEPPSTPCPLRSGHLDNNCARLTLHFNRDHVPQGTIVGALCSGIRSLPATRAVAR 980
QY 34 E-----ASEGPGADEVOV--FAPANALPARS--EAAAVQPVIGISQVRNMSKEKD 81
DB 981 DRLVLCDRASSGASAVEAVSFPARDLPDSSLICGAHAIVAATOR----- 1030
QY 82 LGTLGVLTGMVYIIAIGAGIILG-----YSYKRGKDLKE 118
DB 1031 -GNSSLLAVTEVKEVETVVGSGSTGLVPLCGAFSVLMACVLCVWTRKRRK---- 1085
QY 119 QHDQKVCEREMQRT-----LPLSAFTNPCEIVDEKTVVHTSQTPVDQEGSTP 169
DB 1086 -----ERERSRLPRESANNQWAPLNPRIPIERPGHKDVLVYQCKNFTPPRRADEA 1138
QY 170 LMGQAG 175
DB 1139 LRGPGAG 1144

RESULT 8

US-10-219-248-5
; Sequence 5, Application US/10219248
; Publication No. US200300223688A1
; GENERAL INFORMATION:
; APPLICANT: Sakano, Sei-ji
; APPLICANT: Itoh, Akira
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
; CURRENT APPLICATION NUMBER: US/10/219,248
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US/09/214,278
; PRIOR FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1238
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-219-248-5

Query Match 8.7%; Score 80; DB 9; Length 1238;
Best Local Similarity 22.4%; Pred. No. 13;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;

QY 2 GEAGV--PEKRPC-----EDIRC-----PETTS-----QALPAFTTEIQ 33
DB 947 GEGCAEPPSTPCPLRSGHLDNNCARLTLHFNRDHVPQGTIVGALCSGIRSLPATRAVAR 1006
QY 34 E-----ASEGPGADEVOV--FAPANALPARS--EAAAVQPVIGISQVRNMSKEKD 81
DB 1007 DRLVLCDRASSGASAVEAVSFPARDLPDSSLICGAHAIVAATOR----- 1056
QY 82 LGTLGVLTGMVYIIAIGAGIILG-----YSYKRGKDLKE 118
DB 1057 -GNSSLLAVTEVKEVETVVGSGSTGLVPLCGAFSVLMACVLCVWTRKRRK---- 1111
QY 119 QHDQKVCEREMQRT-----LPLSAFTNPCEIVDEKTVVHTSQTPVDQEGSTP 169
DB 1112 -----ERERSRLPRESANNQWAPLNPRIPIERPGHKDVLVYQCKNFTPPRRADEA 1164
QY 170 LMGQAG 175
DB 1165 LRGPGAG 1170

RESULT 9

US-10-219-247-5
; Sequence 5, Application US/10219247
; Publication No. US20030032781A1
; GENERAL INFORMATION:
; APPLICANT: Sakano, Sei-ji
; APPLICANT: Itoh, Akira
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
; CURRENT APPLICATION NUMBER: US/10/219,247
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US/09/855,722
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/214,278
; PRIOR FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1238
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-219-247-5

Query Match 8.7%; Score 80; DB 9; Length 1238;
Best Local Similarity 22.4%; Pred. No. 13;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;

QY 2 GEAGV--PEKRPC-----EDIRC-----PETTS-----QALPAFTTEIQ 33
DB 947 GEGCAEPPSTPCPLRSGHLDNNCARLTLHFNRDHVPQGTIVGALCSGIRSLPATRAVAR 1006
QY 34 E-----ASEGPGADEVOV--FAPANALPARS--EAAAVQPVIGISQVRNMSKEKD 81
DB 1007 DRLVLCDRASSGASAVEAVSFPARDLPDSSLICGAHAIVAATOR----- 1056
QY 82 LGTLGVLTGMVYIIAIGAGIILG-----YSYKRGKDLKE 118
DB 1057 -GNSSLLAVTEVKEVETVVGSGSTGLVPLCGAFSVLMACVLCVWTRKRRK---- 1111
QY 119 QHDQKVCEREMQRT-----LPLSAFTNPCEIVDEKTVVHTSQTPVDQEGSTP 169
DB 1112 -----ERERSRLPRESANNQWAPLNPRIPIERPGHKDVLVYQCKNFTPPRRADEA 1164
QY 170 LMGQAG 175
DB 1165 LRGPGAG 1170

RESULT 10
US-09-855-722-5

; Sequence 5, Application US/09855722
; Patent No. US20020049306A1
; GENERAL INFORMATION:
; APPLICANT: Sakano, Sei-ji
; APPLICANT: Itoh, Akira
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
; CURRENT APPLICATION NUMBER: US/09/855,722
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/214,278
; PRIOR FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1238
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-855-722-5

Query Match 8.7%; Score 80; DB 10; Length 1238;
Best Local Similarity 22.4%; Pred. No. 13;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;

```

QY 2 GAGV--PEKRC-----EDLRC-----PETTS-----QALPAFTTEIQ 33
DB 947 GCGAEEPPSTPCLPKRGHLDNNCAULTLHFNRDHVPQGTIVGALCSGRISLPATAVAR 1006
QY 34 E-----ASEGPGADEVOV--FAPANALPARS--EAAAVQPVIGISORVMSNEKKD 81
DB 1007 DELVLVLCDRASSASAVEVAFSPARDLPDSSILQGAHAIVMAITQR----- 1056
QY 82 LGTLGVGIGITMMVITIIAIGAGIILG-----YSTKRGDKLKE 118
DB 1057 -GNSLLIATVEVKEVTVGSSSTGLVPLVCGAFSVMLACVULCVMWTRKRRK--- 1111
QY 119 QHDKVCEREMOIT-----LPLSAFTNPCEIYDEKTVVHTSQTPVDPOGSGTP 169
DB 1112 -----ERERSRLPRESANMOMAPLNPTRPIRPGKHVDVLYQCKNTPPPRADEA 1164
QY 170 LMGQAG 175
DB 1165 LPPGAG 1170

```

```

RESULT 11
US-09-944-849-4
; Sequence 4, Application US/09944849
; Patent No. US20020151487A1
; GENERAL INFORMATION:
; APPLICANT: Nickoloff, Brian
; APPLICANT: Miele, Lucio
; TITLE OF INVENTION: METHOD AND REAGENTS FOR EPITHELIAL BARRIER FORMATION AND TREATMEN
; FILE REFERENCE: 212583
; CURRENT APPLICATION NUMBER: US/09/944,849
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,614
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1238
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-944-849-4

```

```

Query Match      8.7%; Score 80; DB 10; Length 1238;
Best Local Similarity 22.4%; Pred. No. 13;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;

```

```

QY 2 GAGV--PEKRC-----EDLRC-----PETTS-----QALPAFTTEIQ 33
DB 947 GCGAEEPPSTPCLPKRGHLDNNCAULTLHFNRDHVPQGTIVGALCSGRISLPATAVAR 1006
QY 34 E-----ASEGPGADEVOV--FAPANALPARS--EAAAVQPVIGISORVMSNEKKD 81
DB 1007 DELVLVLCDRASSASAVEVAFSPARDLPDSSILQGAHAIVMAITQR----- 1056
QY 82 LGTLGVGIGITMMVITIIAIGAGIILG-----YSTKRGDKLKE 118
DB 1057 -GNSLLIATVEVKEVTVGSSSTGLVPLVCGAFSVMLACVULCVMWTRKRRK--- 1111
QY 119 QHDKVCEREMOIT-----LPLSAFTNPCEIYDEKTVVHTSQTPVDPOGSGTP 169
DB 1112 -----ERERSRLPRESANMOMAPLNPTRPIRPGKHVDVLYQCKNTPPPRADEA 1164
QY 170 LMGQAG 175
DB 1165 LPPGAG 1170

```

```

RESULT 12
US-09-921-667-6
; Sequence 6, Application US/09921667
; Patent No. US20020064527A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Mohler, Kendall M.
; APPLICANT: Barone, Dauphine S.
; APPLICANT: Peschon, Jacques J.
; APPLICANT: Kennedy, Mary K.
; APPLICANT: Plummer, John D.
; TITLE OF INVENTION: METHODS FOR TREATING AUTOIMMUNE AND CHRONIC INFLAMMATORY CONDITI
; FILE REFERENCE: 2959-A
; CURRENT APPLICATION NUMBER: US/09/921,667
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/224,079
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 595
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-921-667-6

```

```

Query Match      8.5%; Score 77.5; DB 10; Length 595;
Best Local Similarity 23.6%; Pred. No. 8.9; Indels 43; Gaps 11;
Matches 46; Conservative 32; Mismatches 74;

```

```

QY 4 AGVEKRPCEDLRCPETTSQALPAFT-----TEIOASBEGPGADEVOVAFAPANALPARSEA 59
DB 307 AGETVTRK--QDMAEKDTFEAPPLGTQPDNCNTPENGCAFPAFT-----SPTOSILVDSQA 360
QY 60 AAQVPVIGISORVMSNEKKDGLGVIGITMMVITIIAIGAGIILGYSKRGDKLKEQ 119
DB 361 SKTLF--IPPSAVLSSGKREVLDA--GPVLFVILVIVVVGSSAFL-----L 406
QY 120 HDKVCEREMOIT-----LPLSAFTNPCEIYDEKTVVHTSQTPVDPOGSGTP 166
DB 407 CHRACKRIROKHLICPVOT--SQPLVLVDSRP---RRSTQLRSASVTEPEVBERG 462
QY 167 --STPLMGQAGTPGA 179
DB 463 LMSQPLMETCHSVGA 477

```

```

RESULT 13
US-09-870-759-51
; Sequence 51, Application US/09870759
; Patent No. US2002017551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51
; LENGTH: 1709
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-870-759-51

```

```

Query Match      8.5%; Score 77.5; DB 9; Length 1709;
Best Local Similarity 19.1%; Pred. No. 37; Indels 43; Gaps 6;
Matches 31; Conservative 32; Mismatches 56;
QY 22 SQALPAFTTEI-----QASBEGPGADEVOVAFAPANALPARSEA----- 59
DB 1558 SEPLASITLHLSRVLVASSQPGAPAPPHIVHLASPNALRVDEIALRPSDGEYICSAAN 1617
QY 60 ----AAQVPVIGISORVMSNEKKDGLGVIGITMMVITIIAIGAGIILGYSKRGKD 115
DB 1618 VLGSASTSTYFVALRHRLHFOQ-----LWVLGLVGLVLLILILGAC--YTWR--- 1667

```

QY 116 LKEGHDKVCEREMORITLPLSAFTNPTCEIVDEKTVVHTS 157
 Db 1668 -----RRVCKOSMGENSEVEN-APQKETTOLIDPDATCETS 1702

RESULT 14
 US-09-748-537-14
 ; Sequence 14, Application US/09748537
 ; Patent No. US2002006183A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertin, John
 ; APPLICANT: Chao, Moses V.
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILIE AND USES THERE
 ; FILE REFERENCE: 07314-316001
 ; CURRENT APPLICATION NUMBER: US/09/748,537
 ; PRIOR FILING DATE: 2000-12-26
 ; PRIOR APPLICATION NUMBER: US 09/099,041
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: US 09/019,942
 ; PRIOR FILING DATE: 1998-02-06
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 14
 ; LENGTH: 425
 ; TYPE: PRT
 ; ORGANISM: Rattus rattus
 US-09-748-537-14

Query Match 8.4%; Score 76.5; DB 10; Length 425;
 Best Local Similarity 24.4%; Pred. No. 7.1;
 Matches 42; Conservative 23; Mismatches 90; Indels 17; Gaps 5;

QY 3 EAGVPEKRPCEDLRCPETTSQALPAFTTEIOASGPGADEVQFAPANALPARSEAAAY 62
 Db 175 EROLRECTPWADACEEIPGRWIPRSTPREGSDSTAPSTQEPV-PPEDLVSTVADNV 233
 QY 63 QPVIGISORVRMNSKEKKDLGTLGVLTGMVVIIAIGAGIILGYSYKRGKDLKECHDQ 122
 Db 234 TTVMSQSPVYTR-----GTTDNILPV-YCSILAAVVGVAIVIAFKRMNSCK-QNKQ 284
 QY 123 KVCEREMORITLPLSAFTNPTCEIVDEKTVVHTSOTVPDPOBGSSTPLMGQA 174
 Db 285 GANSR-----PVMOTPPBGEKLSDSGISVDSQSLHDQOTHTQTASGQA 329

RESULT 15
 US-09-821-831-2
 ; Sequence 2, Application US/09821831
 ; Patent No. US20020137188A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bartlett, Perry Francis
 ; APPLICANT: Coulson, Elizabeth Jane
 ; APPLICANT: Fieldew, Katrina
 ; APPLICANT: Baca, Manuel
 ; APPLICANT: Kilpatrick, Trevor
 ; APPLICANT: Surinder, Cheema
 ; TITLE OF INVENTION: Method of Modulating Cell Survival and
 ; FILE REFERENCE: 3206.1001-000
 ; CURRENT APPLICATION NUMBER: US/09/821,831
 ; PRIOR FILING DATE: 2001-03-30
 ; PRIOR APPLICATION NUMBER: PCT/AU99/00860
 ; PRIOR FILING DATE: 1999-10-05
 ; PRIOR APPLICATION NUMBER: AU PQ0701
 ; PRIOR FILING DATE: 1999-06-01
 ; PRIOR APPLICATION NUMBER: AU PP6351
 ; PRIOR FILING DATE: 1998-10-07
 ; PRIOR APPLICATION NUMBER: AU PP6353
 ; PRIOR FILING DATE: 1998-10-06
 ; NUMBER OF SEQ ID NOS: 72
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 425

; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Deduced amino acid sequence
 US-09-821-831-2

Query Match 8.4%; Score 76.5; DB 10; Length 425;
 Best Local Similarity 24.4%; Pred. No. 7.1;
 Matches 42; Conservative 23; Mismatches 90; Indels 17; Gaps 5;
 QY 3 EAGVPEKRPCEDLRCPETTSQALPAFTTEIOASGPGADEVQFAPANALPARSEAAAY 62
 Db 175 EROLRECTPWADACEEIPGRWIPRSTPREGSDSTAPSTQEPV-PPEDLVSTVADNV 233
 QY 63 QPVIGISORVRMNSKEKKDLGTLGVLTGMVVIIAIGAGIILGYSYKRGKDLKECHDQ 122
 Db 234 TTVMSQSPVYTR-----GTTDNILPV-YCSILAAVVGVAIVIAFKRMNSCK-QNKQ 284
 QY 123 KVCEREMORITLPLSAFTNPTCEIVDEKTVVHTSOTVPDPOBGSSTPLMGQA 174
 Db 285 GANSR-----PVMOTPPBGEKLSDSGISVDSQSLHDQOTHTQTASGQA 329

Search completed: April 7, 2003, 09:31:30
 Job time : 14.6528 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:17:08 ; Search time 152.95 Seconds
(without alignments)
754.542 Million cell updates/sec

Title: US-10-057-951-2_COPY_85_263

Perfect score: 916
Sequence: 1 SGKACVPEKPCEDLRCPET.....PVDPRGSLPLMGQATREA 179

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCUS COMB.pep:*

2: /cgn2_6/ptodata/1/paa/US06 COMB.pep:*

3: /cgn2_6/ptodata/1/paa/US07 COMB.pep:*

4: /cgn2_6/ptodata/1/paa/US08 COMB.pep:*

5: /cgn2_6/ptodata/1/paa/US081 COMB.pep:*

6: /cgn2_6/ptodata/1/paa/US082 COMB.pep:*

7: /cgn2_6/ptodata/1/paa/US083 COMB.pep:*

8: /cgn2_6/ptodata/1/paa/US084 COMB.pep:*

9: /cgn2_6/ptodata/1/paa/US085 COMB.pep:*

10: /cgn2_6/ptodata/1/paa/US086 COMB.pep:*

11: /cgn2_6/ptodata/1/paa/US087 COMB.pep:*

12: /cgn2_6/ptodata/1/paa/US088 COMB.pep:*

13: /cgn2_6/ptodata/1/paa/US089 COMB.pep:*

14: /cgn2_6/ptodata/1/paa/US090 COMB.pep:*

15: /cgn2_6/ptodata/1/paa/US091 COMB.pep:*

16: /cgn2_6/ptodata/1/paa/US092 COMB.pep:*

17: /cgn2_6/ptodata/1/paa/US093 COMB.pep:*

18: /cgn2_6/ptodata/1/paa/US094 COMB.pep:*

19: /cgn2_6/ptodata/1/paa/US095 COMB.pep:*

20: /cgn2_6/ptodata/1/paa/US096 COMB.pep:*

21: /cgn2_6/ptodata/1/paa/US097 COMB.pep:*

22: /cgn2_6/ptodata/1/paa/US098 COMB.pep:*

23: /cgn2_6/ptodata/1/paa/US099 COMB.pep:*

24: /cgn2_6/ptodata/1/paa/US100 COMB.pep:*

25: /cgn2_6/ptodata/1/paa/US101 COMB.pep:*

26: /cgn2_6/ptodata/1/paa/US102 COMB.pep:*

27: /cgn2_6/ptodata/1/paa/US60 COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 916 | 100.0 | 250 | 14 | US-09-001-403-64 |
| 2 | 916 | 100.0 | 263 | 1 | PCR-US00-27239-2 |
| 3 | 916 | 100.0 | 263 | 1 | PCR-US02-05301-164 |
| 4 | 916 | 100.0 | 263 | 1 | PCR-US98-10728-2 |
| 5 | 916 | 100.0 | 263 | 14 | US-09-084-491A-2 |
| 6 | 916 | 100.0 | 263 | 15 | US-09-144-889-2 |

| | | | | | | |
|----|-------|-------|------|----|----------------------|-------------------|
| 7 | 916 | 100.0 | 263 | 24 | US-10-057-951-2 | Sequence 2, Appli |
| 8 | 916 | 100.0 | 263 | 25 | US-10-102-704-2 | Sequence 2, Appli |
| 9 | 916 | 100.0 | 263 | 27 | US-60-048-000-2 | Sequence 2, Appli |
| 10 | 913 | 99.7 | 257 | 1 | PCR-US01-14827-8898 | Sequence 8898, Ap |
| 11 | 913 | 99.7 | 263 | 21 | US-09-791-537-39562 | Sequence 39562, A |
| 12 | 913 | 99.7 | 263 | 23 | US-09-927-796-44 | Sequence 44, Appl |
| 13 | 913 | 99.7 | 263 | 26 | US-10-210-951-44 | Sequence 44, Appl |
| 14 | 913 | 99.7 | 263 | 26 | US-10-211-858-44 | Sequence 44, Appl |
| 15 | 913 | 99.7 | 263 | 26 | US-10-211-884-44 | Sequence 44, Appl |
| 16 | 913 | 99.7 | 263 | 27 | US-60-230-435-1560 | Sequence 1560, Ap |
| 17 | 913 | 99.7 | 263 | 27 | US-60-389-987-244 | Sequence 244, App |
| 18 | 913 | 99.7 | 263 | 27 | US-60-412-418-244 | Sequence 244, App |
| 19 | 902 | 98.5 | 263 | 20 | US-09-611-526-3727 | Sequence 3727, Ap |
| 20 | 851.5 | 93.0 | 295 | 1 | PCR-US01-14827-8900 | Sequence 8900, Ap |
| 21 | 807 | 88.1 | 308 | 27 | US-60-207-315-471 | Sequence 471, App |
| 22 | 704.5 | 76.9 | 188 | 15 | US-09-144-889-4 | Sequence 4, Appli |
| 23 | 476 | 52.0 | 178 | 27 | US-60-213-800-237 | Sequence 237, App |
| 24 | 426 | 46.5 | 88 | 22 | US-09-834-366-16200 | Sequence 16200, A |
| 25 | 426 | 46.5 | 88 | 27 | US-60-197-873-16200 | Sequence 16200, A |
| 26 | 322 | 35.2 | 66 | 1 | PCR-US01-00663-31583 | Sequence 31583, A |
| 27 | 322 | 35.2 | 66 | 22 | US-09-864-761-38457 | Sequence 38457, A |
| 28 | 322 | 35.2 | 66 | 25 | US-10-182-993-30642 | Sequence 30642, A |
| 29 | 322 | 35.2 | 66 | 25 | US-10-182-995-24929 | Sequence 24929, A |
| 30 | 322 | 35.2 | 66 | 25 | US-10-182-997-23626 | Sequence 23626, A |
| 31 | 322 | 35.2 | 66 | 26 | US-10-203-134-31343 | Sequence 31343, A |
| 32 | 322 | 35.2 | 66 | 26 | US-10-203-135-30493 | Sequence 30493, A |
| 33 | 322 | 35.2 | 66 | 26 | US-10-203-136-31400 | Sequence 31400, A |
| 34 | 322 | 35.2 | 66 | 26 | US-10-203-137-31583 | Sequence 31583, A |
| 35 | 322 | 35.2 | 66 | 26 | US-10-203-139-30540 | Sequence 30540, A |
| 36 | 275 | 30.0 | 146 | 1 | PCR-US02-05301-237 | Sequence 237, App |
| 37 | 275 | 30.0 | 146 | 1 | PCR-US02-05301-311 | Sequence 311, App |
| 38 | 223 | 24.3 | 126 | 17 | US-09-307-140-979 | Sequence 979, App |
| 39 | 223 | 24.3 | 126 | 22 | US-09-817-076-979 | Sequence 979, App |
| 40 | 135 | 14.7 | 109 | 22 | US-09-834-366-13839 | Sequence 13839, A |
| 41 | 135 | 14.7 | 109 | 27 | US-60-197-873-13839 | Sequence 13839, A |
| 42 | 103 | 11.2 | 117 | 1 | PCR-US01-14827-8897 | Sequence 8897, Ap |
| 43 | 94 | 10.3 | 4752 | 27 | US-60-167-217-14372 | Sequence 14372, A |
| 44 | 94 | 10.3 | 4752 | 27 | US-60-173-464-11654 | Sequence 11654, A |
| 45 | 94 | 10.3 | 4839 | 27 | US-60-191-637-28375 | Sequence 28375, A |

ALIGNMENTS

RESULT 1
US-09-001-403-64
; Sequence 64, Application US/09001403
GENERAL INFORMATION:
APPLICANT: Lai, Preeti
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Au-Young, Janice
APPLICANT: Tang, Y. Tom
APPLICANT: Yue, Henry
APPLICANT: Shah, Purvi
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY PROTEINS
NUMBER OF SEQUENCES: 150
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/001.403

```

; FILING DATE: HERREWITH
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: BILINGS, LUCY J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0455 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 845-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: INODNOT05
; CLONE: 3122252
; US-09-001-403-64

```

```

Query Match          100.0%; Score 916; DB 14; Length 250;
Best Local Similarity 100.0%; Pred. No. 8.6e-90;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SGEAGVPEKRPCEDLCPEPTTSQALPAFTTEIOEASGEGADEVOVFAPANLPAARSEAA 60
DB 72 SGEAGVPEKRPCEDLCPEPTTSQALPAFTTEIOEASGEGADEVOVFAPANLPAARSEAA 131
QY 61 AVQPVIGISQVRMNSKEKKDLGTLGYVLGITMVIITIAIGAGIILGYSKRGKDLKEQH 120
DB 132 AVQPVIGISQVRMNSKEKKDLGTLGYVLGITMVIITIAIGAGIILGYSKRGKDLKEQH 191
QY 121 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGQAGTPGA 179
DB 192 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGQAGTPGA 250

```

```

RESULT 2
PCT-US00-27239-2
; Sequence 2, Application PC/TUS0027239
; GENERAL INFORMATION:
; APPLICANT: HUMAN GENOME SCIENCES, INC.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: P378PCT2
; CURRENT APPLICATION NUMBER: PCT/US00/27239
; CURRENT FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 09/411,977
; PRIOR FILING DATE: 1999-10-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
; PCT-US00-27239-2

```

```

Query Match          100.0%; Score 916; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 9.3e-90;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SGEAGVPEKRPCEDLCPEPTTSQALPAFTTEIOEASGEGADEVOVFAPANLPAARSEAA 60
DB 85 SGEAGVPEKRPCEDLCPEPTTSQALPAFTTEIOEASGEGADEVOVFAPANLPAARSEAA 144
QY 61 AVQPVIGISQVRMNSKEKKDLGTLGYVLGITMVIITIAIGAGIILGYSKRGKDLKEQH 120
DB 145 AVQPVIGISQVRMNSKEKKDLGTLGYVLGITMVIITIAIGAGIILGYSKRGKDLKEQH 204
QY 121 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGQAGTPGA 179
DB 205 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGQAGTPGA 263

```

```

RESULT 3
PCT-US02-05301-164
; Sequence 164, Application PC/TUS0205301
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P5736PCT
; CURRENT APPLICATION NUMBER: PCT/US02/05301
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/304,417
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/270,625
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 164
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
; PCT-US02-05301-164

```

```

Query Match          100.0%; Score 916; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 9.3e-90;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SGEAGVPEKRPCEDLCPEPTTSQALPAFTTEIOEASGEGADEVOVFAPANLPAARSEAA 60
DB 85 SGEAGVPEKRPCEDLCPEPTTSQALPAFTTEIOEASGEGADEVOVFAPANLPAARSEAA 144
QY 61 AVQPVIGISQVRMNSKEKKDLGTLGYVLGITMVIITIAIGAGIILGYSKRGKDLKEQH 120
DB 145 AVQPVIGISQVRMNSKEKKDLGTLGYVLGITMVIITIAIGAGIILGYSKRGKDLKEQH 204
QY 121 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGQAGTPGA 179
DB 205 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGQAGTPGA 263

```

```

RESULT 4
PCT-US98-10728-2
; Sequence 2, Application PC/TUS9810728
; GENERAL INFORMATION:
; APPLICANT: EBNER, REINHARD
; APPLICANT: MOORE, PAUL
; APPLICANT: RUBEN, STEVEN
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/10728
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKS, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: P379PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8439
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids

```

```

QY 1 SGEAGVPEKRPCEDLCPEPTTSQALPAFTTEIOEASGEGADEVOVFAPANLPAARSEAA 60
DB 85 SGEAGVPEKRPCEDLCPEPTTSQALPAFTTEIOEASGEGADEVOVFAPANLPAARSEAA 144
QY 61 AVQPVIGISQVRMNSKEKKDLGTLGYVLGITMVIITIAIGAGIILGYSKRGKDLKEQH 120
DB 145 AVQPVIGISQVRMNSKEKKDLGTLGYVLGITMVIITIAIGAGIILGYSKRGKDLKEQH 204
QY 121 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGQAGTPGA 179
DB 205 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGQAGTPGA 263

```

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US98-10728-2

Query Match 100.0%; Score 916; DB 1; Length 263;

Best Local Similarity 100.0%; Pred. No. 9.3e-90; Indels 0; Gaps 0;

Matches 179; Conservative 0; Mismatches 0;

QY 1 SGEAGVPEKPCEDLRCPEFTTSQALPAFTTEIQEASGPGADEVQVAPANALPARSEAA 60
DB 85 SGEAGVPEKPCEDLRCPEFTTSQALPAFTTEIQEASGPGADEVQVAPANALPARSEAA 144
QY 61 AVQPVIGISQVRVMSKSKKDLGTLGVLGITMVIITIIAGAGIILIGYSYKRGDKLEQH 120
DB 145 AVQPVIGISQVRVMSKSKKDLGTLGVLGITMVIITIIAGAGIILIGYSYKRGDKLEQH 204
QY 121 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGAGTPGA 179
DB 205 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGAGTPGA 263

RESULT 5

US-09-084-491A-2

Sequence 2, Application US/09084491A
GENERAL INFORMATION:
APPLICANT: MOORE, PAUL A.
APPLICANT: RUBEN, STEVEN M.
APPLICANT: BERNER, REINHARD
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/084,491A
FILING DATE: 27-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF378
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-084-491A-2

Query Match 100.0%; Score 916; DB 14; Length 263;

Best Local Similarity 100.0%; Pred. No. 9.3e-90;

Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVPEKPCEDLRCPEFTTSQALPAFTTEIQEASGPGADEVQVAPANALPARSEAA 60
DB 85 SGEAGVPEKPCEDLRCPEFTTSQALPAFTTEIQEASGPGADEVQVAPANALPARSEAA 144
QY 61 AVQPVIGISQVRVMSKSKKDLGTLGVLGITMVIITIIAGAGIILIGYSYKRGDKLEQH 120
DB 145 AVQPVIGISQVRVMSKSKKDLGTLGVLGITMVIITIIAGAGIILIGYSYKRGDKLEQH 204

QY 121 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGAGTPGA 179
DB 205 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGAGTPGA 263

RESULT 6

US-09-144-889-2

Sequence 2, Application US/09144889B
GENERAL INFORMATION:
APPLICANT: Earl F. Albore
APPLICANT: Kristine K. Kikly
TITLE OF INVENTION: KRINGLE1
FILE REFERENCE: GH-70249
CURRENT APPLICATION NUMBER: US/09/144,889B
CURRENT FILING DATE: 1998-09-01
EARLIER APPLICATION NUMBER: 60/056,032
EARLIER FILING DATE: 1997-09-02
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-144-889-2

Query Match 100.0%; Score 916; DB 15; Length 263;

Best Local Similarity 100.0%; Pred. No. 9.3e-90;

Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVPEKPCEDLRCPEFTTSQALPAFTTEIQEASGPGADEVQVAPANALPARSEAA 60
DB 85 SGEAGVPEKPCEDLRCPEFTTSQALPAFTTEIQEASGPGADEVQVAPANALPARSEAA 144
QY 61 AVQPVIGISQVRVMSKSKKDLGTLGVLGITMVIITIIAGAGIILIGYSYKRGDKLEQH 120
DB 145 AVQPVIGISQVRVMSKSKKDLGTLGVLGITMVIITIIAGAGIILIGYSYKRGDKLEQH 204
QY 121 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGAGTPGA 179
DB 205 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGAGTPGA 263

RESULT 7

US-10-057-951-2

Sequence 2, Application US/10057951
GENERAL INFORMATION:
APPLICANT: Moore et al.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
FILE REFERENCE: PF378P1
CURRENT APPLICATION NUMBER: US/10/057,951
CURRENT FILING DATE: 2002-01-29
PRIOR APPLICATION NUMBER: US 09/411,977
PRIOR FILING DATE: 1999-10-04
PRIOR APPLICATION NUMBER: US 09/084,491
PRIOR FILING DATE: 1998-05-27
PRIOR APPLICATION NUMBER: US 60/048,000
PRIOR FILING DATE: 1997-05-28
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-10-057-951-2

Query Match 100.0%; Score 916; DB 24; Length 263;

Best Local Similarity 100.0%; Pred. No. 9.3e-90;

Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVPEKPCEDLRCPEFTTSQALPAFTTEIQEASGPGADEVQVAPANALPARSEAA 60
DB 85 SGEAGVPEKPCEDLRCPEFTTSQALPAFTTEIQEASGPGADEVQVAPANALPARSEAA 144

```
QY 61 AVQPVIGISQVRVNSKEKKDLGTLGYVLGITMMVITIIAGAGIILGYSKRGDLKEQH 120
DB 145 AVQPVIGISQVRVNSKEKKDLGTLGYVLGITMMVITIIAGAGIILGYSKRGDLKEQH 204
QY 121 DQKVCEREMORITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGAGTPGA 179
DB 205 DQKVCEREMORITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGAGTPGA 263

RESULT 8
US-10-102-704-2
; Sequence 2, Application US/10102704
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protein
; FILE REFERENCE: PF378CI
; CURRENT APPLICATION NUMBER: US/10/102,704
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/084,491
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: 60/048,000
; PRIOR FILING DATE: 1997-05-28
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-102-704-2

Query Match 100.0%; Score 916; DB 25; Length 263;
Best Local Similarity 100.0%; Pred. No. 9.3e-90;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVEKRPCEDLRCPEITTSQALPAFTTEIOBASGPGADEVQVAPANALPARSEAA 60
DB 85 SGEAGVEKRPCEDLRCPEITTSQALPAFTTEIOBASGPGADEVQVAPANALPARSEAA 144
QY 61 AVQPVIGISQVRVNSKEKKDLGTLGYVLGITMMVITIIAGAGIILGYSKRGDLKEQH 120
DB 145 AVQPVIGISQVRVNSKEKKDLGTLGYVLGITMMVITIIAGAGIILGYSKRGDLKEQH 204
QY 121 DQKVCEREMORITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGAGTPGA 179
DB 205 DQKVCEREMORITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGAGTPGA 263

RESULT 9
US-60-048-000-2
; Sequence 2, Application US/60048000
; GENERAL INFORMATION:
; APPLICANT: EBNER, REINHARD
; APPLICANT: MOORE, PAUL
; APPLICANT: RUBEN, STEVE
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/048,000
; FILING DATE:
```

```
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: BROOKS, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF379PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-60-048-000-2

Query Match 100.0%; Score 916; DB 27; Length 263;
Best Local Similarity 100.0%; Pred. No. 9.3e-90;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVEKRPCEDLRCPEITTSQALPAFTTEIOBASGPGADEVQVAPANALPARSEAA 60
DB 85 SGEAGVEKRPCEDLRCPEITTSQALPAFTTEIOBASGPGADEVQVAPANALPARSEAA 144
QY 61 AVQPVIGISQVRVNSKEKKDLGTLGYVLGITMMVITIIAGAGIILGYSKRGDLKEQH 120
DB 145 AVQPVIGISQVRVNSKEKKDLGTLGYVLGITMMVITIIAGAGIILGYSKRGDLKEQH 204
QY 121 DQKVCEREMORITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGAGTPGA 179
DB 205 DQKVCEREMORITPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGAGTPGA 263

RESULT 10
PCT-US01-14827-8898
; Sequence 8898, Application PC/TUS0114827
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: PCT/US01/14827
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 8898
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (57)..(78)
; OTHER INFORMATION: KRINGLE DOMAIN SIGNATURE domain identified by eMATRIX,
; OTHER INFORMATION: accession number PR00018C, p-value=5.235e-13, raw score of 14.30
; LOCATION: (25)..(96)
; OTHER INFORMATION: Kringle domain identified by Pfam, accession name kringle, E-
; OTHER INFORMATION: value=1.9e-05, Pfam score of 10.1
PCT-US01-14827-8898

Query Match 99.7%; Score 913; DB 1; Length 257;
Best Local Similarity 99.4%; Pred. No. 1.9e-89;
Matches 178; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVEKRPCEDLRCPEITTSQALPAFTTEIOBASGPGADEVQVAPANALPARSEAA 60
DB 79 SGEAGVEKRPCEDLRCPEITTSQALPAFTTEIOBASGPGADEVQVAPANALPARSEAA 138
QY 61 AVQPVIGISQVRVNSKEKKDLGTLGYVLGITMMVITIIAGAGIILGYSKRGDLKEQH 120
DB 139 AVQPVIGISQVRVNSKEKKDLGTLGYVLGITMMVITIIAGAGIILGYSKRGDLKEQH 198
```


Qy 121 DQKVEREMORTTLPISAFNPTCEIVDEKVVVHTSQTPVDPQEGSTPLMGAGTPGA 179
|||||
Db 199 DQKVEREMORTTLPISAFNPTCEIVDEKVVVHTSQTPVDPQEGSTPLMGAGTPGA 257
|||||

```

RESULT 11
US-09-791-537-39562
; Sequence 39562, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39562
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-39562

```

| | | | | |
|---------------------------|-------|-----------------|-------|------------|
| Query Match | 99.7% | Score 913 | DB 21 | Length 263 |
| Best Local Similarity | 99.4% | Pred. No. 2e-89 | | |
| Matches 178; Conservative | 1 | Mismatches | 0 | Gaps 0 |

| | | | |
|----|-----|--|-----|
| Qy | 1 | SGAGVPEKRPCEDLRCGCTTSQALPATTTIEIOASBEGADEVQVAPAPNALPAREEA | 60 |
| Db | 85 | SGEAGVPEKRPCEDLRCGCTTSQALPATTTIEIOASBEGADEVQVAPAPNALPAREEA | 144 |
| Qy | 61 | AVQPIVIGISQVRNMSKKKDLGTLGYVLGITMVAIIIIAIGAGIIILGYSYKRGDLKEQH | 120 |
| Db | 145 | AVQPIVIGISQVRNMSKKKDLGTLGYVLGITMVAIIIIAIGAGIIILGYSYKRGDLKEQH | 204 |
| Qy | 121 | DQKVEREMQRIITLPLSAFTNPTCEIYDEKVVVHTSGTPVDPDEGSGTPLMGQAGTGA | 179 |
| Db | 205 | DQKVEREMQRIITLPLSAFTNPTCEIYDEKVVVHTSGTPVDPDEGSGTPLMGQAGTGA | 263 |

```

US-09-927-796-44
Sequence 12
Sequence 44, Application US/09927796
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Marsters, Scott A.
APPLICANT: Pan, James
APPLICANT: Picetti, Robert M.
APPLICANT: Roy, Margaret Ann
APPLICANT: Smith, Victoria
APPLICANT: Stone, Donna M.
APPLICANT: Watanabe, Collin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMORS
FILE REFERENCE: P2931R1C1
CURRENT APPLICATION NUMBER: US/09/927, 796
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 60/014699
PRIOR FILING DATE: 1996-04-01
PRIOR APPLICATION NUMBER: 60/026943
PRIOR FILING DATE: 1996-09-23
PRIOR APPLICATION NUMBER: 60/059121
PRIOR FILING DATE: 1997-07-17
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/062037
PRIOR FILING DATE: 1997-10-10

```

| | |
|--------------------------------------|--|
| PRIOR APPLICATION NUMBER: 60/063755 | |
| PRIOR FILING DATE: 1997-10-17 | |
| PRIOR APPLICATION NUMBER: 60/063045 | |
| PRIOR FILING DATE: 1997-10-24 | |
| PRIOR APPLICATION NUMBER: 60/063046 | |
| PRIOR FILING DATE: 1997-10-24 | |
| PRIOR APPLICATION NUMBER: 60/066511 | |
| PRIOR FILING DATE: 1997-11-24 | |
| PRIOR APPLICATION NUMBER: 60/066772 | |
| PRIOR FILING DATE: 1997-11-24 | |
| PRIOR APPLICATION NUMBER: 60/068270 | |
| PRIOR FILING DATE: 1997-12-17 | |
| PRIOR APPLICATION NUMBER: 60/082700 | |
| PRIOR FILING DATE: 1998-04-22 | |
| PRIOR APPLICATION NUMBER: 60/095922 | |
| PRIOR FILING DATE: 1998-08-10 | |
| PRIOR APPLICATION NUMBER: 60/097978 | |
| PRIOR FILING DATE: 1998-08-26 | |
| PRIOR APPLICATION NUMBER: 60/103396 | |
| PRIOR FILING DATE: 1998-10-07 | |
| PRIOR APPLICATION NUMBER: 60/108864 | |
| PRIOR FILING DATE: 1998-11-17 | |
| PRIOR APPLICATION NUMBER: 60/112851 | |
| PRIOR FILING DATE: 1998-12-16 | |
| PRIOR APPLICATION NUMBER: 60/119966 | |
| PRIOR FILING DATE: 1999-02-12 | |
| PRIOR APPLICATION NUMBER: 60/123972 | |
| PRIOR FILING DATE: 1999-03-11 | |
| PRIOR APPLICATION NUMBER: 60/133455 | |
| PRIOR FILING DATE: 1999-05-11 | |
| PRIOR APPLICATION NUMBER: 60/140650 | |
| PRIOR FILING DATE: 1999-06-22 | |
| PRIOR APPLICATION NUMBER: 60/140653 | |
| PRIOR FILING DATE: 1999-06-22 | |
| PRIOR APPLICATION NUMBER: 60/144758 | |
| PRIOR FILING DATE: 1999-07-20 | |
| PRIOR APPLICATION NUMBER: 60/145699 | |
| PRIOR FILING DATE: 1999-07-26 | |
| PRIOR APPLICATION NUMBER: 60/146222 | |
| PRIOR FILING DATE: 1999-07-28 | |
| PRIOR APPLICATION NUMBER: 60/149399 | |
| PRIOR FILING DATE: 1999-08-17 | |
| -PRIOR APPLICATION NUMBER: 60/151688 | |
| PRIOR FILING DATE: 1999-08-31 | |
| PRIOR APPLICATION NUMBER: 60/625328 | |
| PRIOR FILING DATE: 1996-04-01 | |
| PRIOR APPLICATION NUMBER: 60/710802 | |
| PRIOR FILING DATE: 1996-09-23 | |
| PRIOR APPLICATION NUMBER: 60/800699 | |
| PRIOR FILING DATE: 1997-02-14 | |
| PRIOR APPLICATION NUMBER: 60/828683 | |
| PRIOR FILING DATE: 1997-03-31 | |
| PRIOR APPLICATION NUMBER: 60/829270 | |
| PRIOR FILING DATE: 1997-03-31 | |
| PRIOR APPLICATION NUMBER: 60/928066 | |
| PRIOR FILING DATE: 1997-09-11 | |
| PRIOR APPLICATION NUMBER: 60/934494 | |
| PRIOR FILING DATE: 1997-09-19 | |
| PRIOR APPLICATION NUMBER: 60/143066 | |
| PRIOR FILING DATE: 1998-08-28 | |
| PRIOR APPLICATION NUMBER: 60/143707 | |
| PRIOR FILING DATE: 1998-08-28 | |
| PRIOR APPLICATION NUMBER: 60/151889 | |
| PRIOR FILING DATE: 1998-09-11 | |
| PRIOR APPLICATION NUMBER: 60/169104 | |
| PRIOR FILING DATE: 1998-10-05 | |
| PRIOR APPLICATION NUMBER: 60/202088 | |
| PRIOR FILING DATE: 1998-12-08 | |
| PRIOR APPLICATION NUMBER: 60/925331 | |
| PRIOR FILING DATE: 1999-03-03 | |
| PRIOR APPLICATION NUMBER: 60/930403 | |

;; PRIOR FILING DATE: 1999-04-30
;; PRIOR APPLICATION NUMBER: 09/380137
;; PRIOR FILING DATE: 1999-08-25
;; PRIOR APPLICATION NUMBER: 09/380138
;; PRIOR FILING DATE: 1999-08-25
;; PRIOR APPLICATION NUMBER: 09/380139
;; PRIOR FILING DATE: 1999-08-25
;; PRIOR APPLICATION NUMBER: 09/403297
;; PRIOR FILING DATE: 1999-10-18
;; PRIOR APPLICATION NUMBER: 09/423844
;; PRIOR FILING DATE: 1999-11-12
;; PRIOR APPLICATION NUMBER: 09/511133
;; PRIOR FILING DATE: 2000-02-23
;; PRIOR APPLICATION NUMBER: 09/511631
;; PRIOR FILING DATE: 2000-02-23
;; PRIOR APPLICATION NUMBER: 09/664610
;; PRIOR FILING DATE: 2000-09-18
;; PRIOR APPLICATION NUMBER: 09/665350
;; PRIOR FILING DATE: 2000-09-18
;; PRIOR APPLICATION NUMBER: 09/690169
;; PRIOR FILING DATE: 2000-10-16
;; PRIOR APPLICATION NUMBER: 09/690189
;; PRIOR FILING DATE: 2000-10-16
;; PRIOR APPLICATION NUMBER: 09/709238
;; PRIOR FILING DATE: 2000-11-18
;; PRIOR APPLICATION NUMBER: 09/866034
;; PRIOR FILING DATE: 2001-05-25
;; PRIOR APPLICATION NUMBER: 09/872035
;; PRIOR FILING DATE: 2001-06-01
;; PRIOR APPLICATION NUMBER: 09/884733
;; PRIOR FILING DATE: 2001-06-19
;; PRIOR APPLICATION NUMBER: 09/886342
;; PRIOR FILING DATE: 2001-06-19
;; PRIOR APPLICATION NUMBER: 09/866028
;; PRIOR FILING DATE: 2001-08-25
;; PRIOR APPLICATION NUMBER: PCT/US97/05230
;; PRIOR FILING DATE: 1997-03-31
;; PRIOR APPLICATION NUMBER: PCT/US98/19094
;; PRIOR FILING DATE: 1998-09-14
;; PRIOR APPLICATION NUMBER: PCT/US98/19330
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: PCT/US98/21407
;; PRIOR FILING DATE: 1998-10-09
;; PRIOR APPLICATION NUMBER: PCT/US98/25108
;; PRIOR FILING DATE: 1998-12-01
;; PRIOR APPLICATION NUMBER: PCT/US99/05028
;; PRIOR FILING DATE: 1999-03-08
;; PRIOR APPLICATION NUMBER: PCT/US99/12252
;; PRIOR FILING DATE: 1999-06-02
;; PRIOR APPLICATION NUMBER: PCT/US99/20111
;; PRIOR FILING DATE: 1999-09-01
;; PRIOR APPLICATION NUMBER: PCT/US99/21090
;; PRIOR FILING DATE: 1999-09-15
;; PRIOR APPLICATION NUMBER: PCT/US99/28313
;; PRIOR FILING DATE: 1999-11-30
;; PRIOR APPLICATION NUMBER: PCT/US99/28301
;; PRIOR FILING DATE: 1999-12-01
;; PRIOR APPLICATION NUMBER: PCT/US99/28634
;; PRIOR FILING DATE: 1999-12-01
;; PRIOR APPLICATION NUMBER: PCT/US99/28565
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US00/00219
;; PRIOR FILING DATE: 2000-01-05
;; PRIOR APPLICATION NUMBER: PCT/US00/03565
;; PRIOR FILING DATE: 2000-02-11
;; PRIOR APPLICATION NUMBER: PCT/US00/04341
;; PRIOR FILING DATE: 2000-02-18
;; PRIOR APPLICATION NUMBER: PCT/US00/04342
;; PRIOR FILING DATE: 2000-02-18
;; PRIOR APPLICATION NUMBER: PCT/US00/04414
;; PRIOR FILING DATE: 2000-02-22
;; PRIOR APPLICATION NUMBER: PCT/US00/05841
;; PRIOR FILING DATE: 2000-03-02

;; PRIOR APPLICATION NUMBER: PCT/US00/06884
;; PRIOR FILING DATE: 2000-03-15
;; PRIOR APPLICATION NUMBER: PCT/US00/08439
;; PRIOR FILING DATE: 2000-03-30
;; PRIOR APPLICATION NUMBER: PCT/US00/13705
;; PRIOR FILING DATE: 2000-05-17
;; PRIOR APPLICATION NUMBER: PCT/US00/14941
;; PRIOR FILING DATE: 2000-05-30
;; PRIOR APPLICATION NUMBER: PCT/US00/15264
;; PRIOR FILING DATE: 2000-06-02
;; PRIOR APPLICATION NUMBER: PCT/US00/20710
;; PRIOR FILING DATE: 2000-07-28
;; PRIOR APPLICATION NUMBER: PCT/US00/32678
;; PRIOR FILING DATE: 2000-12-01
;; PRIOR APPLICATION NUMBER: PCT/US01/17800
;; PRIOR FILING DATE: 2001-06-01
;; PRIOR APPLICATION NUMBER: PCT/US01/19692
;; PRIOR FILING DATE: 2001-06-20
;; PRIOR APPLICATION NUMBER: PCT/US01/21066
;; PRIOR FILING DATE: 2001-06-29
;; PRIOR APPLICATION NUMBER: PCT/US01/21735
;; PRIOR FILING DATE: 2001-07-09
;; NUMBER OF SEQ ID NOS: 258
;; SEQ ID NO 44
;; LENGTH: 263

Query Match 99.7%; Score 913; DB 23; Length 263;
Best Local Similarity 99.4%; Pred. No. 2e-89; Indels 0; Gaps 0;
Matches 178; Conservative 1; Mismatches 0;

QY 1 SGEAGVPEKRCEDLRCPETTSQALPAFTTEIQEASGSGPADEYQVAPANALPARSEAA 60
DB 85 SGEAGVPEKRCEDLRCPETTSQALPAFTTEIQEASGSGPADEYQVAPANALPARSEAA 144
QY 61 AVQPVIGISQVRNNSKEKDLGTLGYLGTVMVVIILAGITLLGYSKRGDLKEQH 120
DB 145 AVQPVIGISQVRNNSKEKDLGTLGYLGTVMVVIILAGITLLGYSKRGDLKEQH 204
QY 121 DQKVCEREMORITLPLSAFTNPTCEIVDEKTVVHTSOTPDPOEGSTPLMGAGTPGA 179
DB 205 DQKVCEREMORITLPLSAFTNPTCEIVDEKTVVHTSOTPDPOEGSTPLMGAGTPGA 263

RESULT 13
US-10-210-951-44
;; Sequence 44, Application US/10210951
;; GENERAL INFORMATION:
;; APPLICANT: Ashkenazi, Avi J.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Hillan, Kenneth J.
;; APPLICANT: Marsters, Scot A.
;; APPLICANT: Pan, James
;; APPLICANT: Pilti, Robert M.
;; APPLICANT: Roy, Margaret Ann
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stone, Donna M.
;; APPLICANT: Watanebe, Colin K.
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
;; FILE REFERENCE: P2931R1C1
;; CURRENT APPLICATION NUMBER: US/10/210,951
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: 60/014699
;; PRIOR FILING DATE: 1996-04-01
;; PRIOR APPLICATION NUMBER: 60/026943
;; PRIOR FILING DATE: 1996-09-23
;; PRIOR APPLICATION NUMBER: 60/059121
;; PRIOR FILING DATE: 1997-07-17
;; PRIOR APPLICATION NUMBER: 60/059352
;; PRIOR FILING DATE: 1997-09-19
;; PRIOR APPLICATION NUMBER: 60/062037

```

; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-951-44
```

```

Query Match          99.7%; Score 913; DB 26; Length 263;
Best Local Similarity 99.4%; Pred. No. 2e-89; 0; Indels 0; Gaps 0;
Matches 178; Conservative 1; Mismatches 0;
```

```

Qy 1 SGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIQEASBEGGADDEVQVFAPANALPARSEAA 60
Db 85 SGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIQEASBEGGADDEVQVFAPANALPARSEAA 144
Qy 61 AVQPVIGISQVRVMSKKEKDLGTGYVIGITMVIYIIAGIIGIIGSYKRGKDLKEQH 120
Db 145 AVQPVIGISQVRVMSKKEKDLGTGYVIGITMVIYIIAGIIGIIGSYKRGKDLKEQH 204
Qy 121 DQKVCEREMQRTITPLSAFTNPTCEIVDEKTVVHTSOTPVDPQEGSTPLMGAGTPGA 179
Db 205 DQKVCEREMQRTITPLSAFTNPTCEIVDEKTVVHTSOTPVDPQEGSTPLMGAGTPGA 263
```

```

RESULT 14
US-10-211-858-44
; Sequence 44, Application US/10211858
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scott A.
; APPLICANT: Pan, James
; APPLICANT: Pitti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Matanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/211,858
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
```

```

; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-858-44
```

```

Query Match          99.7%; Score 913; DB 26; Length 263;
Best Local Similarity 99.4%; Pred. No. 2e-89; 0; Indels 0; Gaps 0;
Matches 178; Conservative 1; Mismatches 0;
```

```

Qy 1 SGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIQEASBEGGADDEVQVFAPANALPARSEAA 60
Db 85 SGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIQEASBEGGADDEVQVFAPANALPARSEAA 144
Qy 61 AVQPVIGISQVRVMSKKEKDLGTGYVIGITMVIYIIAGIIGIIGSYKRGKDLKEQH 120
Db 145 AVQPVIGISQVRVMSKKEKDLGTGYVIGITMVIYIIAGIIGIIGSYKRGKDLKEQH 204
Qy 121 DQKVCEREMQRTITPLSAFTNPTCEIVDEKTVVHTSOTPVDPQEGSTPLMGAGTPGA 179
Db 205 DQKVCEREMQRTITPLSAFTNPTCEIVDEKTVVHTSOTPVDPQEGSTPLMGAGTPGA 263
```

```

RESULT 15
US-10-211-884-44
; Sequence 44, Application US/10211884
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scott A.
; APPLICANT: Pan, James
; APPLICANT: Pitti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Matanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/211,884
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
```

; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-884-44

Query Match 99.7%; Score 913; DB 26; Length 263;
Best Local Similarity 99.4%; Pred. No. 2e-89; 0; Indels 0; Gaps 0;
Matches 178; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| Qy | 1 | SGEAGVPEKRPCEDLRCPEETTSQALPAPFTTEIQEASEGPGADEVOVFAPANALPARSEAA | 60 |
| Db | 85 | SGEAGVPEKRPCEDLRCPEETTSQALPAPFTTEIQEASEGPGADEVOVFAPANALPARSEAA | 144 |
| Qy | 61 | AVQPVIGISORVAMNSKEKKDLGTGVLGITMVTIIAIGAGIILGYSYKRGKDLKEQH | 120 |
| Db | 145 | AVQPVIGISORVAMNSKEKKDLGTGVLGITMVTIIAIGAGIILGYSYKRGKDLKEQH | 204 |
| Qy | 121 | DQVCEREMORTLPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGAGTPGA | 179 |
| Db | 205 | DQVCEREMORTLPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGAGTPGA | 263 |

Search completed: April 7, 2003, 09:29:28
Job time : 153.95 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:18:38 ; Search time 26.7942 Seconds
(without alignments)
1010.317 Million cell updates/sec

Title: US-10-057-951-2_COPY_85_263

Perfect score: 916
Sequence: 1 SGEAGVPEKRPCEDLRCPEF.....PVDPOGSTRPLMGAGTGA 179

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 679521 seqs, 151232488 residues

Total number of hits satisfying chosen parameters: 679521

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--|
| 1 | 913 | 99.7 | 263 | 6 | US-10-218-140-6002 Sequence 6002, App |
| 2 | 732.5 | 80.0 | 264 | 6 | US-10-144-779-488 Sequence 488, App |
| 3 | 95 | 10.4 | 163 | 1 | PCT-US02-332727-117039 Sequence 17039, A |
| 4 | 95 | 10.4 | 163 | 5 | US-09-978-825-17039 Sequence 17039, A |
| 5 | 95 | 10.4 | 163 | 6 | US-10-057-998-17039 Sequence 17039, A |
| 6 | 87.5 | 9.6 | 850 | 6 | US-10-369-493-10111 Sequence 10111, A |
| 7 | 86.5 | 9.4 | 507 | 6 | US-10-289-762-174 Sequence 174, App |
| 8 | 86.5 | 9.4 | 820 | 6 | US-10-282-122A-54852 Sequence 54852, A |
| 9 | 80.5 | 8.8 | 769 | 5 | US-09-701-271A-4 Sequence 4, Appli |
| 10 | 80 | 8.7 | 345 | 5 | US-09-724-676-63523 Sequence 63523, A |
| 11 | 80 | 8.7 | 345 | 5 | US-09-724-676A-63523 Sequence 63523, A |
| 12 | 80 | 8.7 | 1226 | 5 | US-09-724-676-63525 Sequence 63525, A |
| 13 | 80 | 8.7 | 1226 | 5 | US-09-724-676A-63525 Sequence 63525, A |
| 14 | 80 | 8.7 | 1264 | 5 | US-09-724-676-63524 Sequence 63524, A |
| 15 | 80 | 8.7 | 1264 | 5 | US-09-724-676A-63524 Sequence 63524, A |
| 16 | 79.5 | 8.7 | 769 | 5 | US-09-701-271A-6 Sequence 6, Appli |
| 17 | 79.5 | 8.7 | 769 | 6 | US-10-320-800-10 Sequence 10, Appli |
| 18 | 78.5 | 8.6 | 2824 | 6 | US-10-369-493-5166 Sequence 5166, App |
| 19 | 77.5 | 8.5 | 452 | 6 | US-10-282-122A-70552 Sequence 70552, App |
| 20 | 77.5 | 8.5 | 463 | 6 | US-10-092-411A-3973 Sequence 3973, App |
| 21 | 77.5 | 8.5 | 977 | 7 | US-60-440-068-526 Sequence 526, App |
| 22 | 77.5 | 8.5 | 1709 | 5 | US-09-751-708A-51 Sequence 51, Appl |
| 23 | 77.5 | 8.5 | 1709 | 5 | US-09-949-016-10503 Sequence 10503, A |
| 24 | 77.5 | 8.5 | 1709 | 7 | US-60-440-068-168 Sequence 168, App |
| 25 | 77.5 | 8.5 | 1709 | 7 | US-60-453-135-14575 Sequence 14575, A |
| 26 | 77.5 | 8.5 | 1709 | 7 | US-60-453-050-14575 Sequence 14575, A |

| | | | | | |
|----|------|-----|-----|---|--|
| 27 | 77 | 8.4 | 769 | 5 | US-09-701-271A-2 Sequence 2, Appli |
| 28 | 76 | 8.3 | 490 | 6 | US-10-369-493-21753 Sequence 21753, A |
| 29 | 76 | 8.3 | 755 | 6 | US-10-156-761-11835 Sequence 11835, A |
| 30 | 75.5 | 8.2 | 219 | 6 | US-10-092-411A-5651 Sequence 5651, App |
| 31 | 75.5 | 8.2 | 276 | 6 | US-10-203-138A-14692 Sequence 14692, A |
| 32 | 75.5 | 8.2 | 422 | 1 | PCT-US02-24310-134 Sequence 134, App |
| 33 | 75 | 8.2 | 595 | 5 | US-09-949-016-6048 Sequence 6048, App |
| 34 | 75 | 8.2 | 595 | 5 | US-09-826-212A-9 Sequence 9, Appl |
| 35 | 75 | 8.2 | 595 | 7 | US-60-443-566-4128 Sequence 4128, App |
| 36 | 75 | 8.2 | 595 | 7 | US-60-452-680-20991 Sequence 20991, A |
| 37 | 75 | 8.2 | 595 | 7 | US-60-453-135-12718 Sequence 12718, A |
| 38 | 75 | 8.2 | 595 | 7 | US-60-453-135-12718 Sequence 12718, A |
| 39 | 75 | 8.2 | 642 | 5 | US-09-949-016-8043 Sequence 8043, App |
| 40 | 75 | 8.2 | 774 | 5 | US-09-905-846A-69 Sequence 69, Appl |
| 41 | 75 | 8.2 | 774 | 5 | US-09-948-429C-69 Sequence 69, Appl |
| 42 | 75 | 8.2 | 788 | 6 | US-10-156-761-14497 Sequence 14497, A |
| 43 | 75 | 8.2 | 950 | 6 | US-10-366-683-29012 Sequence 29012, A |
| 44 | 74.5 | 8.1 | 340 | 6 | US-10-282-122A-48299 Sequence 48299, A |
| 45 | 74.5 | 8.1 | 630 | 7 | US-60-452-680-23441 Sequence 23441, A |

ALIGNMENTS

```
RESULT 1
US-10-218-140-6002
; Sequence 6002, Application US/10218140
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYPEPTIDES
; FILE REFERENCE: 15966-543 CON
; CURRENT APPLICATION NUMBER: US/01218, 140
; PRIOR FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/540, 763
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/127, 728
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/127, 636
; PRIOR FILING DATE: 1999-04-02
; PRIOR APPLICATION NUMBER: 60/127, 607
; NUMBER OF SEQ ID NOS: 6322
; SOFTWARE: Curator Version 1.0
; SEQ ID NO 6002
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-218-140-6002
Query Match 99.7%; Score 913, DB 6, Length 263;
Best Local Similarity 99.4%; Pred. No. 7.3e-81;
Matches 178; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVPEKRPCEDLRCPEFSSQALPAFTTRIOASBEGADVEQVAPNAPLARSEEA 60
DB 85 SGEAGVPEKRPCEDLRCPEFSSQALPAFTTRIOASBEGADVEQVAPNAPLARSEEA 144
QY 61 AVDPVIGISQRYVNMSEKEDLGTLYGVYIGITVMYIIAIGAGIILGYKRGKDLKEOH 120
DB 145 AVDPVIGISQRYVNMSEKEDLGTLYGVYIGITVMYIIAIGAGIILGYKRGKDLKEOH 204
QY 121 DQVVCEREMQRTPLPSAFTTPTCEIVDEKVVVHTSQTPTDPOGSTRPLMGAGTGA 179
DB 205 DQVVCEREMQRTPLPSAFTTPTCEIVDEKVVVHTSQTPTDPOGSTRPLMGAGTGA 263

RESULT 2
US-10-144-779-488
; Sequence 488, Application US/10144779
; GENERAL INFORMATION:
; APPLICANT: SUBRAMANIAN, Mani et al.
```

```
; TITLE OF INVENTION: MOUSE ORTHOLOGS OF HUMAN DISEASE GENES,  
; TITLE OF INVENTION: PROTEINS ENCODED BY THESE MOUSE GENES, AND USES THEREOF  
; FILE REFERENCE: CL001235  
; CURRENT APPLICATION NUMBER: US/10/144,779  
; CURRENT FILING DATE: 2002-05-15  
; NUMBER OF SEQ ID NOS: 864  
; SOFTWARE: FastSeq for windows version 4.0  
; SEQ ID NO 488  
; LENGTH: 264  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-144-779-488
```

```
Query Match 80.0%; Score 732.5; DB 6; Length 264;  
Best Local Similarity 79.4%; Pred. No. 2.8e-63;  
Matches 143; Conservative 13; Mismatches 23; Indels 1; Gaps 1;
```

```
QY 1 SGEAGVEKRPCELDLCPETTSQA-LPAFTTEIOEASSEGGADEVQVAPANALPARESEA 59  
DB 85 SSETGVKRPCEDEVSCPETTSQAPPPSAMLEEKSGAPGKEKQVFPANALPARESEA 144  
QY 60 AAVQVIGISQVRNMSKEKKDGLGYLGYITMVIITAIAGIILGYSKRGKDLKEQ 119  
DB 145 AEVQVIGISQVRNMSKEKKDGLGYLGYITMVIITAIAGIIVGYTGRGKDLKEQ 204  
QY 120 HDQKVCEREMORTLPISAFNPTCEIVDEKVVVHTSQTVPDQSGTPLMGAKTGEA 179  
DB 205 HEKACEREMORTLPISAFNPTCEIVDEKVVVHTSQTVPDQSGTLLTGQNGTGEA 264
```

```
RESULT 3  
PCT-US02-32727-17039  
; Sequence 17039, Application PC/TUS0232727
```

```
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer  
; APPLICANT: Skelky, Yasir  
; APPLICANT: Persing, David  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Maisonneuve, Jean Francois  
; APPLICANT: Zhang, Yanni  
; APPLICANT: Wang, Siqing  
; APPLICANT: Jen, Shyian  
; APPLICANT: Lodes, Michael  
; APPLICANT: Benson, Darin  
; APPLICANT: Jones, Robert  
; APPLICANT: Carter, Derrick  
; APPLICANT: Barth, Brenda  
; APPLICANT: Douglas, John  
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
; FILE REFERENCE: 210121.514C1  
; CURRENT APPLICATION NUMBER: PCT/US02/32727  
; CURRENT FILING DATE: 2002-10-11  
; NUMBER OF SEQ ID NOS: 30992  
; SEQ ID NO 17039  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Propionibacterium acnes  
PCT-US02-32727-17039
```

```
Query Match 10.4%; Score 95; DB 1; Length 163;  
Best Local Similarity 26.1%; Pred. No. 0.18;  
Matches 40; Conservative 30; Mismatches 47; Indels 36; Gaps 9;
```

```
QY 49 PANALPARESEAAVQVPI-GISQVRNMSKEKKDGLGYLGYITMVIITAIAG-AGIIL 106  
DB 18 PQQARPAARPTSLRPVLRALPQVR-----QQQWGRLGFAI--LIVVMLAAGLAGLIV 69  
QY 107 -----GYSYKRGKDLKE-----QHDQKVCEREMORTLP-----LSAFNPTCEIV 147  
DB 78 LNTTIOAQSMQIAQETRNINLVLQHQAVALAAEVDHLRGQNVQDEAKKLGMRPNVGSYI 129  
QY 148 DEKTVVHTSQTVPDQSGT-----GST--PLMGQ 173
```

```
DB 130 DLRTGKVIQTKVDGKEVPGVIGETAKPEVQ 162
```

```
RESULT 4  
US-09-978-825-17039  
; Sequence 17039, Application US/09978825
```

```
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer  
; APPLICANT: Skelky, Yasir  
; APPLICANT: Persing, David  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Maisonneuve, Jean Francois  
; APPLICANT: Zhang, Yanni  
; APPLICANT: Wang, Siqing  
; APPLICANT: Jen, Shyian  
; APPLICANT: Lodes, Michael  
; APPLICANT: Benson, Darin  
; APPLICANT: Jones, Robert  
; APPLICANT: Carter, Derrick  
; APPLICANT: Barth, Brenda  
; APPLICANT: Douglas, John  
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
; FILE REFERENCE: 210121.514C1  
; CURRENT APPLICATION NUMBER: US/09/978,825  
; CURRENT FILING DATE: 2003-01-29  
; NUMBER OF SEQ ID NOS: 30992  
; SEQ ID NO 17039  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Propionibacterium acnes  
US-09-978-825-17039
```

```
Query Match 10.4%; Score 95; DB 5; Length 163;  
Best Local Similarity 26.1%; Pred. No. 0.18;  
Matches 40; Conservative 30; Mismatches 47; Indels 36; Gaps 9;
```

```
QY 49 PANALPARESEAAVQVPI-GISQVRNMSKEKKDGLGYLGYITMVIITAIAG-AGIIL 106  
DB 18 PQQARPAARPTSLRPVLRALPQVR-----QQQWGRLGFAI--LIVVMLAAGLAGLIV 69  
QY 107 -----GYSYKRGKDLKE-----QHDQKVCEREMORTLP-----LSAFNPTCEIV 147  
DB 70 LNTTIOAQSMQIAQETRNINLVLQHQAVALAAEVDHLRGQNVLQDEAKKLGMRPNVGSYI 129  
QY 148 DEKTVVHTSQTVPDQSGT-----GST--PLMGQ 173  
DB 130 DLRTGKVIQTKVDGKEVPGVIGETAKPEVQ 162
```

```
RESULT 5  
US-10-057-498-17039  
; Sequence 17039, Application US/10057498
```

```
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer  
; APPLICANT: Skelky, Yasir  
; APPLICANT: Persing, David  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Maisonneuve, Jean Francois  
; APPLICANT: Zhang, Yanni  
; APPLICANT: Wang, Siqing  
; APPLICANT: Jen, Shyian  
; APPLICANT: Lodes, Michael  
; APPLICANT: Benson, Darin  
; APPLICANT: Jones, Robert  
; APPLICANT: Carter, Derrick  
; APPLICANT: Barth, Brenda  
; APPLICANT: Douglas, John  
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
; FILE REFERENCE: 210121.514  
; CURRENT APPLICATION NUMBER: US/10/057,498  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 29212  
; SEQ ID NO 17039  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Propionibacterium acnes  
US-10-057-498-17039
```

```
Query Match 10.4%; Score 95; DB 6; Length 163;  
Best Local Similarity 26.1%; Pred. No. 0.18;  
Matches 40; Conservative 30; Mismatches 47; Indels 36; Gaps 9;  
QY 49 PANALPARESEAAVQVPI-GISQVRNMSKEKKDGLGYLGYITMVIITAIAG-AGIIL 106
```

```
Db 16 POQARTAAARPTSLRPVLRALPQVR-----QQOMKRLGAI---LIVMLAAGLAGLIV 69
Qy 107 -----GYSYKRGKDKKE-----QHDQKVCEREMORTLP-----LSAFTNPTCEIV 147
Db 70 LNTTIOAOMQAOETRYNINVLQHOAVLAEVDHLRGPNLOEAKKIGMBPNFGSYI 129
Qy 148 DEKTVVHTSQTVPDPOE-----GST--PLMGQ 173
Db 130 DLRGKVIQGTQTKVDGKVPVIGSETAKPEVIGQ 162

RESULT 6
US-10-369-493-10111
; Sequence 10111, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 10111
; LENGTH: 850
; TYPE: PRF
; ORGANISM: magnetite-containing magnetic coccus
US-10-369-493-10111

Query Match
Best Local Similarity 22.2%; Score 87.5; DB 6; Length 850;
Matches 41; Conservative 18; Mismatches 73; Indels 53; Gaps 4;

Qy 11 PCEDLRCPETTSQALPAFTT-----IQASGPGADEVQVFAPANLPPARSEAAVQPV- 65
Db 72 PAAPVAPAPAPAPVAVVAVPVAQAPAPAPVVEEMVAKPLPAPAAAPASAPVR 131
Qy 66 -----IGISQVRMNSKEKKDGLTGYVLGITMWTYIIAIGAGTILGYSYR 112
Db 132 LMQPVEEPRKRSKQGRREMARKTEDVVS-----KR 165
Qy 113 GMDLKEQHOKVCEREMORTLPASAFTNPTCEIVDEKTVVHTSQTVPDPOGSLPLMG 172
Db 166 LMQLELRQKKEDARKEAEVALAK-----KEKPVAAATAAAAEVVAAGRTPRD 216
Qy 173 QAGTP 177
Db 217 SAGEP 221

RESULT 7
US-10-289-762-174
; Sequence 174, Application US/10289762
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevet
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 174
; LENGTH: 507
; TYPE: PRF
; ORGANISM: Chlamydia pneumoniae
; NAME/KEY: SITE
```

```
; LOCATION: 1...507
; OTHER INFORMATION: Xaa=unknown or other
US-10-289-762-174

Query Match
Best Local Similarity 24.1%; Score 86.5; DB 6; Length 507;
Matches 38; Conservative 28; Mismatches 49; Indels 43; Gaps 9;

Qy 20 TTSQALPFTTEIODE-----ASEGPGADEVQVFAPANLPPARSEAA-VQPVIGIS 69
Db 235 IQEGSLFAFTTRLDITLLGVSFLVIAPEHDPDISV-----SEQRDEVTAYVDESIRKS 288
Qy 70 QRVNNS-KKKKDLGTGY-----VLGITMVV-----IIAIGAGTILGYSYKRGKDKKEQ 119
Db 289 ERDRISVKTQVFTGNVAKHPITGNLLPVMISDVVLGVTGVVWGV-----PA 339
Qy 120 HQQKVCEREMORTLPASAFTNPTCEIVDEKTVVHTS 157
Db 340 HDER--DREFAEM-----FSLPIHEVIDNGVCIHSN 369

RESULT 8
US-10-282-122A-54852
; Sequence 54852, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITPA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SORTNAME: PatentIn version 3.1
; SEQ ID NO 54852
; LENGTH: 820
; TYPE: PRF
; ORGANISM: Chlamydia pneumoniae
US-10-282-122A-54852

Query Match
Best Local Similarity 9.4%; Score 86.5; DB 6; Length 820;
Matches 38; Conservative 28; Mismatches 49; Indels 43; Gaps 9;
```

```

QY 20 TTSCALPAFTTELOE-----ASEPGADEVQAFAPALPARSEAAA-VQPVGIS 69
Db 235 TQSGLEAFTTRIDLTGLGSPVIAPEHPDLDSIV-----SEQRDEYTAVOESLRKS 288
QY 70 QRRNMS-KEKDLGLTG---VLGITMY---IIIGAGIILGYSYRQDLKEQ 119
Db 289 EPRRISVTKTGTVFGNVAKPIKINLPLWISDYVLGIGYGVYMGV-----PA 339
QY 120 HDQKVCERQRIITLPLAFNPTCEIYDEKTVYHRS 157
Db 340 HDER--DREPAEM-----FSLPIHEVIDDQGVCIHNS 369

```

```

RESULT 9
US-09-701-271A-4
Sequence 4, Application US/09701271A
GENERAL INFORMATION:
APPLICANT: Ruelle, Jean-Louis
APPLICANT: Tommassen, Johannes Petrus Maria
TITLE OF INVENTION: Neisseria Meningitidis Antigenic
TITLE OF INVENTION: Polypeptides, Corresponding Polynucleotides and Protective
TITLE OF INVENTION: Antibodies
FILE REFERENCE: BM45323
CURRENT APPLICATION NUMBER: US/09/701,271A
CURRENT FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: PCT/EP99/03603
PRIOR FILING DATE: 1999-05-26
PRIOR APPLICATION NUMBER: 9811260.0
PRIOR FILING DATE: 1998-05-26
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 769
TYPE: PR1
ORGANISM: Neisseria meningitidis
US-09-701-271A-4

```

```

Query March 8.8%; Score 80.5; DB 5; Length 765;
Best Local Similarity 24.4%; Pred. No. 33;
Matches 49; Conservative 26; Mismatches 69; Indels 57; Gaps 11.

Cy 3 EAGPEKRPEDLRCPETTS--QALPAFTTEIGAS-----EGGADENVQFAPA-- 50
Db 150 QAAAPSTKSAVSQKPTPAKQAAAPFTSEVSISAPSPAKQAAASAKQTTAPAKQ 209
Cy 51 -NALPASEAAAVQPVIGISQVRVMSKERKDLGTLGVYLGITMMVTIIAIGAGIIGIS 109
Db 210 QTAAPAAQQAAPAKQOTNIDFR-----KQGNAG-----IIEIAA---LGF 248
Cy 110 YKRGKDLKEHDOKVCEEREMORTITLPLSAFTPT-----CEIVDEKTVV--VHTSQTPVD 162
Db 249 GQ--FDISQGH-----HIVTLKQHTLPTTLQRLSDVADPKPVPQVLTAKRLNND 257
Cy 163 PQ-----EGSNPLMGAGATPG 178
Db 298 TQLITTAGNWLNVKSNABG 318

RESULT 10
US-09-724-676-63523
; Sequence 63523, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 63523
; LENGTH: 345
; TYPE: PRT
;

```

```

! ORGANISM: Homo sapiens
US-09-TG2A-676-63523

Query Match      8.7%; Score 60; DB 5; Length 345;
Best local similarity 22.4%; Pred. No. 13;
Matches    55; Conservative   20; Mismatches    77; Indels   94; Gaps     11./

QY       2  GGAGV--PEKRP-----EDLR-----PETTS-----QAAPFTTEIQ  33
           |||         |::|||          :|||
Db        54  GECAEPPPTPLPSGHLDNNCARLTLHRNDHVQGTTGAICSGIRSLPATAVAR  113

QY       34  E-----ASEGPDAEVQ-FEPANALPARS-EAAAVPYIGISQRVRMSKEKO  81
           ::            ::|||::|::|::|::|::|::|::|::|::|
Db        114 DRLVLCDRASSGVAVEVAVSFPARDLSLLGGAAHAIVAAITGR---  163

QY       82  LGTGLVGLTMMVVIIIAIGAIIIIG-----YSYRGKDLE  118
           ::|||::|::|::|::|::|::|::|::|::|
Db        164 -GNSSLILATEVEKVTVVTGSSSTGLLVPILCAFSVLMACVCLCVMTTRERRK----  218

QY       119 OHODKCVERMORIT-----LPISAFTNPCEIYDEKTVMHTSQTPVPDQGSTP  169
           |||         |||::|::|::|::|::|::|::|
Db        219 -----ERRRSLPRRESANNOMAPLNIPINPLERRGKHDLVLYQCOKNTPPREBRADA  271

QY       170 LMGQAG 175
           |||         |||
Db        272 LPGPAG 277
```

```

RESULT 11
US-09-724-676A-63523
: Sequence 63523, Application US/09724676A
: GENERAL INFORMATION:
: APPLICANT: CompuGen LTD
: TITLE OF INVENTION: Variants of alternative splicing
: FILE REFERENCE: 129181.4 CompuGen
: CURRENT APPLICATION NUMBER: US/09/724.676A
: CURRENT FILING DATE: 2000-11-28
: NUMBER OF SEQ ID NOS: 97222
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 63523
: LENGTH: 345
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-724-676A-63523

```

```

Query Match: 8 7%; Score 80; DB 5; Length 345;
Best Local Similarity 22.4%; Fred. No. 13;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11.

QY      2  GEAGV--PEKRPC-----EDLRC-----PETTS-----QALPATTETQ 33
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      54  GECAGEPETPCPLPSGHLDNNCARLTLFNRDHPVQGTVGALCSGIRSLPATRAVAR 113

QY      34  E-----ASTGPADEVOY-FAPAAALPARS--EAAAVQPIGSGCVRNKNSREKD 81
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      114  DRLVLICDRASSGASAVEAVSFSPPARDLPSSLSIOGAHAIVAAITQ----- 163

QY      82  LGTLGVLTITWVITIIAIGAILIG-----YSYRGKOLAE 118
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      164  GNSLLLAATVEKVEIVTVTGSSFTGLLVPLCGAFSVLMLACVILCVMWTRRRRK---- 218

QY      119  QHDKVCEERMCQIT-----LPLSAFNPTCEIYDEKTVVHNSQPVDPQSGSTP 169
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      219  -----ERRRSRLPRESANNQMAPLNPITNPILERPGCHKVDLYQCKNTFPPPRADA 271

QY      170  LMGQAG 175
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      272  LRGPGAG 277

RESULT 12
US-09-724-676-63525
; Sequence 63525, Application US/09724676

```



```

; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 63525
; LENGTH: 1226
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (561)..(561)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (565)..(565)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-724-676-63525
```

```

Query Match
Best Local Similarity 22.4%; Score 80; DB 5; Length 1226;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;
```

```

QY 2 GEGAV--PEKPPC-----EDLRC-----PETTS-----QALPATTETIQ 33
DB 935 GEGAEPPSTPCLPKSGHLNNCARLTLHFNRDHVPQGTIVGALCSGIRSLPTRAVAR 994
QY 34 E-----ASEGPGADEVOV--FAPANALPARS--EAAAVQPIGISCVRNMSKEKD 81
DB 995 DRLVLVLCRRASSGASAVEVAVSFSPARDLPDSSLIGAAHAIVAAITQR----- 1044
QY 82 LGTLGVVGIITMWTIIAIGAGIILG-----YSYRGKDLKE 118
DB 1045 -GNSSLLAVTEVKEVETVVGSSSTGLVPLVLCGAFSVLMACVLCVWTRKRRK--- 1099
QY 119 QHDQKVCEREMORT-----LPLSAFTNPCEIVDEKTVVHTSQTVPVDQEGSTP 169
DB 1100 -----ERERSRLPREBSANNQMAPLNPINPIERPGHKDVLVYOCKKFTPPRADEA 1152
QY 170 LMGQAG 175
DB 1153 LPPGAG 1158
```

```

RESULT 13
US-09-724-676A-63525
; Sequence 63525, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 63525
; LENGTH: 1226
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (561)..(561)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (565)..(565)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-724-676A-63525
```

```

Query Match 8.7%; Score 80; DB 5; Length 1226;
```

```

Best Local Similarity 22.4%; Pred. No. 66;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;

QY 2 GEGAV--PEKPPC-----EDLRC-----PETTS-----QALPATTETIQ 33
DB 935 GEGAEPPSTPCLPKSGHLNNCARLTLHFNRDHVPQGTIVGALCSGIRSLPTRAVAR 994
QY 34 E-----ASEGPGADEVOV--FAPANALPARS--EAAAVQPIGISCVRNMSKEKD 81
DB 995 DRLVLVLCRRASSGASAVEVAVSFSPARDLPDSSLIGAAHAIVAAITQR----- 1044
QY 82 LGTLGVVGIITMWTIIAIGAGIILG-----YSYRGKDLKE 118
DB 1045 -GNSSLLAVTEVKEVETVVGSSSTGLVPLVLCGAFSVLMACVLCVWTRKRRK--- 1099
QY 119 QHDQKVCEREMORT-----LPLSAFTNPCEIVDEKTVVHTSQTVPVDQEGSTP 169
DB 1100 -----ERERSRLPREBSANNQMAPLNPINPIERPGHKDVLVYOCKKFTPPRADEA 1152
QY 170 LMGQAG 175
DB 1153 LPPGAG 1158
```

```

RESULT 14
US-09-724-676-63524
; Sequence 63524, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 63524
; LENGTH: 1264
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (599)..(599)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (603)..(603)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-724-676-63524
```

```

Query Match 8.7%; Score 80; DB 5; Length 1264;
Best Local Similarity 22.4%; Pred. No. 69;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;
```

```

QY 2 GEGAV--PEKPPC-----EDLRC-----PETTS-----QALPATTETIQ 33
DB 973 GEGAEPPSTPCLPKSGHLNNCARLTLHFNRDHVPQGTIVGALCSGIRSLPTRAVAR 1032
QY 34 E-----ASEGPGADEVOV--FAPANALPARS--EAAAVQPIGISCVRNMSKEKD 81
DB 1033 DRLVLVLCRRASSGASAVEVAVSFSPARDLPDSSLIGAAHAIVAAITQR----- 1082
QY 82 LGTLGVVGIITMWTIIAIGAGIILG-----YSYRGKDLKE 118
DB 1083 -GNSSLLAVTEVKEVETVVGSSSTGLVPLVLCGAFSVLMACVLCVWTRKRRK--- 1137
QY 119 QHDQKVCEREMORT-----LPLSAFTNPCEIVDEKTVVHTSQTVPVDQEGSTP 169
DB 1138 -----ERERSRLPREBSANNQMAPLNPINPIERPGHKDVLVYOCKKFTPPRADEA 1190
QY 170 LMGQAG 175
DB 1191 LPPGAG 1196
```

RESULT 15
US-09-724-676A-63524
; Sequence 63524, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 63524
; LENGTH: 1264
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (599)..(599)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (603)..(603)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-724-676A-63524

Query Match 8.7%; Score 80; DB 5; Length 1264;

Best Local Similarity 22.4%; Pred. No. 69; Mismatches 77; Indels 94; Gaps 11;

Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;

QY 2 GEAGV--PEKRPC-----EDLRC-----PETTS-----QALPATTETIQ 33
DB 973 GEGCAEPPSTCLPRSGHLNNCARLTLEHNDHVPQGTWGAICSGIRSLPATRAVAR 1032
QY 34 E-----ASEGPADEVQV-FAPANALPARS--EAAAVQPYIGISQREVMNSXKKD 81
DB 1033 DRLLVLLCDRASSGASAVEVAVSFPARDLPDSLLIOGAHAIVAAITQR----- 1082
QY 82 LGTGVVLGITMMVILIIAGITLG-----YSYRKGDLKE 118
DB 1083 -GNSSLIAVTEVKVETVTGSGTGLVPLGAPSVMLACVLCVWMTTRRRK---- 1137
QY 119 QHOKVCEREMQRI-----LPSAFNPTCEIVDEKTVVYHNSQTPVDQEGSTP 169
DB 1138 -----ERRSRLPRESANNQWAPLPIRNPRIERPGHKDVLQCKNFTPPRADDA 1190
QY 170 LMGQAG 175
DB 1191 LPGPAG 1196

Search completed: April 7, 2003, 09:30:51
Job time : 29.7942 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:28 ; Search time 17.8628 Seconds
(without alignments)
294.842 Million cell updates/sec

Title: US-10-057-951-2_COPY_85_263

Perfect score: 916
Sequence: 1 SGEAGVPEKRPEDLRCPET.....PVDQEGSTPLMGQAGTPGA 179

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: Issued Patents AA.*
2: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/6CTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 916 | 100.0 | 263 | 4 | US-09-411-977-2 |
| 2 | 80 | 8.7 | 1212 | 4 | US-09-214-278-3 |
| 3 | 80 | 8.7 | 1238 | 4 | US-09-214-278-5 |
| 4 | 77.5 | 8.5 | 463 | 4 | US-09-134-001C-3973 |
| 5 | 77.5 | 8.5 | 595 | 1 | US-08-225-989-2 |
| 6 | 77.5 | 8.5 | 595 | 1 | US-08-570-923-2 |
| 7 | 77.5 | 8.5 | 595 | 1 | US-08-580-014-2 |
| 8 | 77.5 | 8.5 | 595 | 4 | US-09-079-785-2 |
| 9 | 75.5 | 8.2 | 219 | 4 | US-09-134-001C-5651 |
| 10 | 75 | 8.2 | 595 | 2 | US-08-232-087A-2 |
| 11 | 75 | 8.2 | 595 | 4 | US-09-006-353A-9 |
| 12 | 75 | 8.2 | 595 | 4 | US-09-573-986-9 |
| 13 | 74.5 | 8.1 | 587 | 1 | US-08-398-008A-2 |
| 14 | 74.5 | 8.1 | 587 | 1 | US-08-893-333-2 |
| 15 | 74 | 8.1 | 1148 | 4 | US-08-883-046-4 |
| 16 | 72.5 | 7.9 | 225 | 4 | US-09-134-001C-3215 |
| 17 | 72.5 | 7.9 | 400 | 1 | US-08-351-473B-5 |
| 18 | 72.5 | 7.9 | 400 | 4 | US-08-450-962-4 |
| 19 | 72.5 | 7.9 | 400 | 4 | US-08-450-962-6 |
| 20 | 72 | 7.9 | 1780 | 1 | US-08-769-309A-5 |
| 21 | 72 | 7.9 | 1780 | 1 | US-08-994-570-5 |
| 22 | 71.5 | 7.8 | 707 | 4 | US-09-228-986-80 |
| 23 | 71 | 7.8 | 652 | 4 | US-09-751-305-2 |
| 24 | 70.5 | 7.7 | 1182 | 4 | US-09-287-354-6 |
| 25 | 69.5 | 7.6 | 1025 | 3 | US-08-304-309-4 |
| 26 | 69.5 | 7.6 | 1025 | 2 | US-08-991-942-4 |
| 27 | 69 | 7.5 | 1025 | 2 | US-08-304-309-2 |

| | | | | | | |
|----|------|-----|-------|---|---------------------|--------------------|
| 28 | 69 | 7.5 | 1025 | 3 | US-08-991-942-2 | Sequence 2, Appl1 |
| 29 | 69 | 7.5 | 1025 | 4 | US-09-138-103-2 | Sequence 2, Appl1 |
| 30 | 69 | 7.5 | 1025 | 5 | PCT-US95-04567-4 | Sequence 4, Appl1 |
| 31 | 68.5 | 7.5 | 673 | 2 | US-08-455-073A-6 | Sequence 6, Appl1 |
| 32 | 68.5 | 7.5 | 683 | 2 | US-08-477-196A-17 | Sequence 17, Appl1 |
| 33 | 68 | 7.4 | 1257 | 3 | US-08-611-729A-8 | Sequence 8, Appl1 |
| 34 | 67.5 | 7.4 | 683 | 1 | US-07-878-960-2 | Sequence 2, Appl1 |
| 35 | 67 | 7.3 | 271 | 4 | US-09-414-436-1 | Sequence 1, Appl1 |
| 36 | 67 | 7.3 | 388 | 1 | US-08-087-772A-2 | Sequence 2, Appl1 |
| 37 | 67 | 7.3 | 458 | 4 | US-09-134-001C-4663 | Sequence 4, Appl1 |
| 38 | 67 | 7.3 | 531 | 4 | US-09-134-001C-4920 | Sequence 4663, Ap |
| 39 | 66.5 | 7.3 | 400 | 1 | US-08-351-473B-4 | Sequence 4920, Ap |
| 40 | 66.5 | 7.3 | 1248 | 4 | US-08-882-046-6 | Sequence 4, Appl1 |
| 41 | 66.5 | 7.3 | 10182 | 4 | US-09-134-001C-3159 | Sequence 6, Appl1 |
| 42 | 66 | 7.2 | 279 | 4 | US-08-339-214-24 | Sequence 3159, Ap |
| 43 | 66 | 7.2 | 279 | 4 | US-08-339-214-32 | Sequence 24, Appl1 |
| 44 | 66 | 7.2 | 427 | 4 | US-09-086-483A-4 | Sequence 32, Appl1 |
| 45 | 66 | 7.2 | 427 | 4 | US-09-041-886-2 | Sequence 4, Appl1 |

ALIGNMENTS

```
RESULT 1
US-09-411-977-2
; Sequence 2, Application US/09411977
; Patent No. 6372473
; GENERAL INFORMATION:
; APPLICANT: Moore, Paul A.
; APPLICANT: Ruben, Steven M.
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: pe378P1
; CURRENT APPLICATION NUMBER: US/09/411,977
; CURRENT FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: 09/084,491
; EARLIER FILING DATE: 1998-05-27
; EARLIER APPLICATION NUMBER: 60/048,000
; EARLIER FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-411-977-2

Query Match          100.0%; Score 916; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 8.1e-98;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVPEKRPEDLRCPETTSQALPAFTTBIOASGPGADEQVAPANALPARSEAA 60
DB 85 SGEAGVPEKRPEDLRCPETTSQALPAFTTBIOASGPGADEQVAPANALPARSEAA 144

QY 61 AVQPIGISQVRNMSKKKDLGTLGYVLGITMVIIAIGAGIILGYSYKRRDKLEQH 120
DB 145 AVQPIGISQVRNMSKKKDLGTLGYVLGITMVIIAIGAGIILGYSYKRRDKLEQH 204

QY 121 DQKVEREMORTPLSFTNPTCEIYDEKVYVHTSOTVDPQEGSTPLMGAGTPGA 179
DB 205 DQKVEREMORTPLSFTNPTCEIYDEKVYVHTSOTVDPQEGSTPLMGAGTPGA 263

RESULT 2
US-09-214-278-3
; Sequence 3, Application US/09214278
; Patent No. 6291210
; GENERAL INFORMATION:
; APPLICANT: Sakano, Seiji
; APPLICANT: Itoh, Akira
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
```

;; CURRENT APPLICATION NUMBER: US/09/214,278
;; CURRENT FILING DATE: 1999-01-26
;; NUMBER OF SEQ ID NOS: 32
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 3
;; LENGTH: 1212
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-214-278-3

Query Match 8.7%; Score 80; DB 4; Length 1212;
Best Local Similarity 22.4%; Pred.No. 3.4;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;

QY 2 GEAGV--PEKRPC-----EDLRC-----PETTS-----QALPFTTEIQ 33
DB 921 GEGCAEPPSTPCLPRSGHLDNNCARLTHFNRDHVPQGTGALCSGIRSLPATRAVAR 980
QY 34 E-----ASBGPQADEYQV-FAPANMLPARS--EAAAVPVGISQRYVMNSKEKD 81
DB 981 DRLIVLCDBRASSGASAVEAVSFSPARDLPSSLIQGAHAIVAATQR----- 1030
QY 82 LGTLGVVLGITMMVITIAIGAGIILG-----YSYKRGKDLKE 118
DB 1031 -GNSSLILAVTEYKVEVTVGSSSTGLVPLCGAFSVLMACVLCVWMTKRK----- 1085
QY 119 QHDKVCEREMQRT-----LPLSAFTNPTCEIYDEKTVVHTSOTPVDPQEGSTP 169
DB 1086 -----ERRSRLLPREESANNOMAPLNPRIERPGRGHKQVLYCKKNFTPPPRADEA 1138
QY 170 LMGQAG 175
DB 1139 LPEPAG 1144

RESULT 3
US-09-214-278-5
;; Sequence 5, Application US/09214278
;; Patent No. 6291210
;; GENERAL INFORMATION:
;; APPLICANT: Itoh, Akira
;; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
;; FILE REFERENCE: KP-8576
;; CURRENT APPLICATION NUMBER: US/09/214,278
;; CURRENT FILING DATE: 1999-01-26
;; NUMBER OF SEQ ID NOS: 32
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 5
;; LENGTH: 1238
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-214-278-5

Query Match 8.7%; Score 80; DB 4; Length 1238;
Best Local Similarity 22.4%; Pred.No. 3.5;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;

QY 2 GEAGV--PEKRPC-----EDLRC-----PETTS-----QALPFTTEIQ 33
DB 947 GEGCAEPPSTPCLPRSGHLDNNCARLTHFNRDHVPQGTGALCSGIRSLPATRAVAR 1006
QY 34 E-----ASBGPQADEYQV-FAPANMLPARS--EAAAVPVGISQRYVMNSKEKD 81
DB 1007 DRLIVLCDBRASSGASAVEAVSFSPARDLPSSLIQGAHAIVAATQR----- 1056
QY 82 LGTLGVVLGITMMVITIAIGAGIILG-----YSYKRGKDLKE 118
DB 1057 -GNSSLILAVTEYKVEVTVGSSSTGLVPLCGAFSVLMACVLCVWMTKRK----- 1111
QY 119 QHDKVCEREMQRT-----LPLSAFTNPTCEIYDEKTVVHTSOTPVDPQEGSTP 169
DB 1112 -----ERRSRLLPREESANNOMAPLNPRIERPGRGHKQVLYCKKNFTPPPRADEA 1164

QY 170 LMGQAG 175
DB 1165 LPEPAG 1170

RESULT 4
US-09-134-001C-3973
;; Sequence 3973, Application US/09134001C
;; Patent No. 6380370
;; GENERAL INFORMATION:
;; APPLICANT: Lynn Doucette-Stamm et al
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;; FILE REFERENCE: GTC-007
;; CURRENT APPLICATION NUMBER: US/09/134,001C
;; CURRENT FILING DATE: 1998-08-13
;; PRIOR APPLICATION NUMBER: US 60/064,964
;; PRIOR FILING DATE: 1997-11-08
;; PRIOR FILING DATE: 1997-08-14
;; NUMBER OF SEQ ID NOS: 5674
;; SEQ ID NO 3973
;; LENGTH: 463
;; TYPE: PRT
;; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3973

Query Match 8.5%; Score 77.5; DB 4; Length 463;
Best Local Similarity 31.1%; Pred.No. 1.6;
Matches 28; Conservative 14; Mismatches 37; Indels 11; Gaps 3;

QY 66 IGISQR--VRMSKKEKDLGTLGVVLGITMMVITIA-IGAGIILGYKRGKDLKEHQD 122
DB 116 IGIRERQILMLNDRDPTNSGIVKLIEIVRTIFIEFIALLAFFYRDNPLK----- 170
QY 123 KVCEREMQRTPLPSAFTNPTCEIYDEKTV 152
DB 171 ---NALMGIFVSVATNGIGLIDIGESLV 197

RESULT 5
US-08-225-989-2
;; Sequence 2, Application US/08225989
;; Patent No. 5480981
;; GENERAL INFORMATION:
;; APPLICANT: Goodwin, Raymond G.
;; APPLICANT: Smith, Craig A.
;; APPLICANT: Armitage, Richard J.
;; APPLICANT: Grusec, Hans-Jurgen
;; TITLE OF INVENTION: No. 5480981el Cytokine That Binds CD30
;; NUMBER OF SEQUENCES: 23
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Kathryn A. Seese, Immunex Corporation
;; STREET: 51 University Street
;; CITY: Seattle
;; STATE: Washington
;; COUNTRY: USA
;; ZIP: 98101
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: Apple Macintosh
;; OPERATING SYSTEM: Apple 7.1
;; SOFTWARE: Microsoft Word, Version 5.1a
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/225,989
;; FILING DATE: 12 APRIL 1994
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/966,775
;; FILING DATE: 27-OCT-1992
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 907,224
FILING DATE: 01-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 899,660
FILING DATE: 15-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 892,459
FILING DATE: 02-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 889,717
FILING DATE: 26-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Seeee, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2804-E
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-225-989-2

Query Match 8.5%; Score 77.5; DB 1; Length 595;
Best Local Similarity 23.6%; Pred. No. 2.3;
Matches 46; Conservative 32; Mismatches 74; Indels 43; Gaps 11;

QY 4 AGVPEKPCEDLRCEETTSQALPAFT---TEIQASGPGADVGVFAPNALPARESA 59
DB 307 AGETVTKR-QDMAKQDTFFBAPLGTQPCNPFPENGSAFST-----SPTQSLVDSQA 360
QY 60 AAVQPVIGISQVRNMSKEDLGTGLGVLTGMVYIIAIGAGIIGSYVKRQDLKEQ 119
DB 361 SKTLP-IPTSAPVALSSTGKPVDA-GPVLFWYILVIVVVGSSAFV-----L 406
QY 120 HDQKVCEREMQR---ITPLSAFTNPCEIVDEKTVVHTSQT-----PYDPQEG 166
DB 407 CHRRACKRKIRQKHLHCVVQT-SQPLFLVDSRP---RRSSTQLRSGASVTEPVAEREG 462
QY 167 --STPLMGQAGTPGA 179
DB 463 LMSQPLMETCHSVGA 477

RESULT 6
US-08-570-923-2
Sequence 2, Application US/08570923
Patent No. 5677430
GENERAL INFORMATION:
APPLICANT: Goodwin, Raymond G.
APPLICANT: Smith, Craig A.
APPLICANT: Armitage, Richard J.
APPLICANT: Gruss, Hans-Jurgen
TITLE OF INVENTION: No. 5677430el Cytokine That Binds CD30
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kathryn A. Seeee, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple 7.1
SOFTWARE: Microsoft Word, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,923

FILING DATE: 12-DEC-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/225,989
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: US 07/966,775
FILING DATE: 27-OCT-1992
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 907,224
FILING DATE: 01-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 899,660
FILING DATE: 15-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 892,459
FILING DATE: 02-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 889,717
FILING DATE: 26-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Seeee, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2804-E
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-923-2

Query Match 8.5%; Score 77.5; DB 1; Length 595;
Best Local Similarity 23.6%; Pred. No. 2.3;
Matches 46; Conservative 32; Mismatches 74; Indels 43; Gaps 11;

QY 4 AGVPEKPCEDLRCEETTSQALPAFT---TEIQASGPGADVGVFAPNALPARESA 59
DB 307 AGETVTKR-QDMAKQDTFFBAPLGTQPCNPFPENGSAFST-----SPTQSLVDSQA 360
QY 60 AAVQPVIGISQVRNMSKEDLGTGLGVLTGMVYIIAIGAGIIGSYVKRQDLKEQ 119
DB 361 SKTLP-IPTSAPVALSSTGKPVDA-GPVLFWYILVIVVVGSSAFV-----L 406
QY 120 HDQKVCEREMQR---ITPLSAFTNPCEIVDEKTVVHTSQT-----PYDPQEG 166
DB 407 CHRRACKRKIRQKHLHCVVQT-SQPLFLVDSRP---RRSSTQLRSGASVTEPVAEREG 462
QY 167 --STPLMGQAGTPGA 179
DB 463 LMSQPLMETCHSVGA 477

RESULT 7
US-08-580-014-2
Sequence 2, Application US/08580014
Patent No. 5753203
GENERAL INFORMATION:
APPLICANT: Goodwin, Raymond G.
APPLICANT: Smith, Craig A.
APPLICANT: Armitage, Richard J.
APPLICANT: Gruss, Hans-Jurgen
TITLE OF INVENTION: No. 5753203el Cytokine That Binds CD30
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kathryn A. Seeee, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington

```

COUNTRY: USA
ZIP: 96101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple 7.1
SOFTWARE: Microsoft Word, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/580.014
FILING DATE: 20-DEC-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/225.989
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: US 07/966.775
FILING DATE: 27-OCT-1992
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 907.224
FILING DATE: 01-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 899.660
FILING DATE: 15-JUN-1992
PRICE APPLICATION DATA:
APPLICATION NUMBER: US 892.459
FILING DATE: 02-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 889.717
FILING DATE: 26-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Seese, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2804-E
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-580-014-2

Query Match      8.5%; Score 77.5; DB 1; Length 595;
Best Local Similarity 23.6%; Pred. No. 2.3;
Matches 46; Conservative 32; Mismatches 74; Indels 43; Gaps 11.

OY    4 AGVBRKRPEDIRCEPETSQALPAFT-----TETDASEGPGADEVOVFAPANALPARSEA 59
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db    307 AGETVTKP-QDMAEDDTTFEAPLGTGPDPGNPTPENGEAPAST-----SPTQSLLVDSEA 160

OY    60 AAQVPITGISQKVRWNSKEKKDGTGYVLGITWWVIIIAIGAILGVSYRGDKLEQ 119
       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    361 SKLP-IPIPSAVLASSTCKRFLDA-GPLFPVILIVLVVVGSATL-----L 406

OY    120 HDQVEREMQR---ITPLSAFINPCGIUDEKTVVHTSQ-----PVDPDEG 166
       ::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    407 CHRRACKRKRIKRLHLCLYPQT-SQPTELVDSP--RRSSTQLMSGASVTPEABERG 462

OY    167 --STPLMGAGTPGA 179
       |||::|||::|||
Db    463 LMSQPLMERCHSVGA 477

RESULT 8
US-09-079-785-2
Sequence 2, Application US/09079785
Patent No. 6143869
GENERAL INFORMATION:
APPLICANT: Goodwin, Raymond G.
APPLICANT: Smith, Craig A.
```

```

APPLICANT: Amritage, Richard J.
APPLICANT: Grues, Hans-Jürgen
TITLE OF INVENTION: No. 6143869e1 Cytokine That Binds CD30
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kathryn A. Seese, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple 7.1
SOFTWARE: Microsoft Word, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/079,785
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/225,989
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: US 07/966,775
FILING DATE: 27-OCT-1992

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 907,224
FILING DATE: 01-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 899,660
FILING DATE: 15-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 892,459
FILING DATE: 02-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 889,717
FILING DATE: 26-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Seese, Kathryn A
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2804-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)233-0644
TELEX: 756822

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-079-785-2

Query Match      8.5%; Score 77.5; DB 4; Length 595;
Best Local Similarity 23.6%; Pred. No.2.3; Indels 43; Gaps 11.;
Matches 46; Conservative 32; Mismatches

QY    4 AGVKKRPCEIDRCPTTSOALPAFT---TEIQASGFGADEVQVFAPNALPASEA 59
       : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    307 AGGTATVR-QWAKARDTFEAPPLIGDPCNPTENGEAFAST----SPTQSLLVDQA 360
       : : : : : : : : : : : : : : : : : : : : : : : : : :
QY    60 AAQPVTGISQRVMNSKEKSDIGTLGVLTITMVIIIAIGAGIILGYSKRKDLKEQ 119
       : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    361 SKLP-LIPSAFVALISSTGRVLDA-GPVLFMWIVLVVVVGSSAFYL-----L 406
       : : : : : : : : : : : : : : : : : : : : : : : : : :
QY    120 HDQKVEMRMR---ITLPSAPTNPCEIVEDEKTVMVHVSOT-----PYDPQEG 166
       : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    407 CHRRARRKRIRQGKLHCYPVQT-SGPXKELVDSRP---RSSQTQRSAGSVTEPPVAERG 462
       : : : : : : : : : : : : : : : : : : : : : : : : : :
QY    167 -STPLMGAGTPEGA 179
       ||| |
Db    463 LMSQPLMETCHSGVA 477

```

```

RESULT 9
US-09-134-001C-5651
; Sequence 5651, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5651
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5651

Query Match          8.2%; Score 75.5; DB 4; Length 219;
Best Local Similarity 26.7%; Pred. No. 0.88;
Matches 32; Conservative 22; Mismatches 33; Indels 33; Gaps 7;

QY 75 NSKEKDL-GTLGVLTGTMVILIAAGIILGYSY-----KRG-----113
DB 7 NREVNNSGIGSKFSAIVLILIGLAFGI--YSFVDSKKGNERLSDKTTQOEKKD 63

QY 114 -KDLKEOHDOVCVCRE---MORTLPLSAFTNPTCEIYDEKTVVHTSQT-PPV--DPQEG 166
DB 64 DKDKKKDKKSKVEKKNKNTQQTQVQVQTQQTQQTQV--OTPPRPPTTQTPVKNPQTG 121

RESULT 10
US-08-232-087A-2
; Sequence 2, Application US/08232087A
; Patent No. 5866372
; GENERAL INFORMATION:
; APPLICANT: Stein, Harald
; APPLICANT: D Rkop, Horst
; APPLICANT: Latza, Ute
; TITLE OF INVENTION: Lymphoid CD30-Antigen
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Birch, Stewart, Kolaesch & Birch, LLP
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,087A
; FILING DATE: 08-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 756-103P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 203-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 595 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-232-087A-2

Query Match          8.2%; Score 75; DB 2; Length 595;
Best Local Similarity 23.1%; Pred. No. 4.4;
Matches 43; Conservative 31; Mismatches 70; Indels 42; Gaps 10;

QY 13 EDLRCPTTSQALPAPT-----TEIQASGPGADVQVFAPAMLPARSAAVQPVITGI 68
DB 315 QDMVKQTTTPEAPLGTOPDCNTPENGEPAST-----SPQSILVDSQASKTLF-IPT 368

QY 69 SQRVMSKKEKDLGTLGVLTGTMVILIAAGIILGYSYKKGKDLKEQHDQVCERE 128
DB 369 SAPVALSTGKPLVLA-GVLFVWILVTVVVGSAFL-----LCHRRACRRK 415

QY 129 MQR---ITPLSLAFTNPTCEIYDEKTVVHTSQT-----PVPDQEG--STPLMGQ 173
DB 416 IROKLHLCEYVQT-SQPKELVDSRP---RRSTQLRSGASVTEPVAEERGLMSQPLMET 471

QY 174 AGTPGA 179
DB 472 CHSVGA 477

RESULT 11
US-09-006-353A-9
; Sequence 9, Application US/09006353A
; Patent No. 6261801
; GENERAL INFORMATION:
; APPLICANT: WEI, YING-PEI
; APPLICANT: YU, CHIO-LIANG
; APPLICANT: GENTZ, REINER
; APPLICANT: RUBEN, STEVEN
; TITLE OF INVENTION: TUMOR NECROSIS FACTOR RECEPTOR 5
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/006,353A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF341
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 595 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-006-353A-9

Query Match          8.2%; Score 75; DB 4; Length 595;
Best Local Similarity 23.1%; Pred. No. 4.4;
Matches 43; Conservative 31; Mismatches 70; Indels 42; Gaps 10;

```

QY 13 EDLNCPEPTTSQALPAFT---TEIOEASGPGADEVQVAPAPNALPARSEAAVQPVIGI 68
DB 315 QDMAEKDTTPEAPPLGTQPCNPTPENGEPAST-----SPTQSLVDSQASKTLR-1PT 368
QY 69 SQRPMNSKEKKDLGTGLGYLGTMMVITIIAGAGIIGYSYKRGKDLKEQHDQVCERE 126
DB 369 SAPVALSSTGKPVILDA-GPVLFWITLVVVVVGSSAFL-----LCHRRACKRK 415
QY 129 MOR---ITLPLSAFTNPTCEIVDEKTVVHTSQT-----PVDPQEG--STPLMGQ 173
DB 416 IRQKLHLCPYVQT-SQPKLELVDSRP---RRSSTQLRSGASVTEPAERGIAMSGPLMET 471
QY 174 AGTRGA 179
DB 472 CHSVGA 477

RESULT 12

US-09-573-986-9
Sequence 9, Application US/09573986
Patent No. 6455040
GENERAL INFORMATION:
APPLICANT: Wei, Ying-Fei
APPLICANT: Ni, Jian
APPLICANT: Gentz, Reiner
APPLICANT: Ruben, Steven
TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
FILE REFERENCE: 1488,1280004
CURRENT APPLICATION NUMBER: US/09/573,986
CURRENT FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patentn Ver. 2.1
SEQ ID NO 9
LENGTH: 595
TYPE: PRT
ORGANISM: Homo sapiens
US-09-573-986-9

Query Match

Best Local Similarity 23.4%; Score 75; DB 4; Length 595;
Matches 43; Conservative 31; Mismatches 70; Indels 42; Gaps 10;

QY 13 EDLNCPEPTTSQALPAFT---TEIOEASGPGADEVQVAPAPNALPARSEAAVQPVIGI 68
DB 315 QDMAEKDTTPEAPPLGTQPCNPTPENGEPAST-----SPTQSLVDSQASKTLR-1PT 368
QY 69 SQRPMNSKEKKDLGTGLGYLGTMMVITIIAGAGIIGYSYKRGKDLKEQHDQVCERE 126
DB 369 SAPVALSSTGKPVILDA-GPVLFWITLVVVVVGSSAFL-----LCHRRACKRK 415
QY 129 MOR---ITLPLSAFTNPTCEIVDEKTVVHTSQT-----PVDPQEG--STPLMGQ 173
DB 416 IRQKLHLCPYVQT-SQPKLELVDSRP---RRSSTQLRSGASVTEPAERGIAMSGPLMET 471
QY 174 AGTRGA 179
DB 472 CHSVGA 477

RESULT 13

US-08-398-008A-2
Sequence 2, Application US/08398008A
Patent No. 5665588

GENERAL INFORMATION:
APPLICANT: Kornbluth, Jacki
TITLE OF INVENTION: DNA Encoding Natural Killer Lytic Associated
NUMBER OF INVENTION: Protein
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gilbreth & Adler, P.C.
STREET: 8011 Candle Lane
CITY: Houston

STATE: Texas
COUNTRY: USA
ZIP: 77071
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage
COMPUTER: MACINTOSH IICX
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/398,008A
FILING DATE: March 2, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/126,501
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Dr. Benjamin Aaron
REGISTRATION NUMBER: 35,423
REFERENCE/DOCKET NUMBER: D5705CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 777-2321
TELEFAX: (713) 777-6908
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 587
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: Protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-398-008A-2

Query Match

Best Local Similarity 8.1%; Score 74.5; DB 1; Length 587;
Matches 38; Conservative 17; Mismatches 39; Indels 57; Gaps 8;

QY 52 ALPRAEAAVQPVIGISQEV-----RMSSEKKDLGTGLGYLGTMMVITIIAGAGIIGYSY 97
DB 367 ALPAM-----VIGIPVYGRKHRSYGRKTSKHNLAITG---GVLSVYASPI 415
QY 98 ---IAIGAGIIGYSY-----KRGDLK---EQHQKVCEREMQKIT 133
DB 416 AAVSGIGVPIPLMAYVGVVPISLCRGGGVSTANGKGVKLEPDEDGPITVADNR-- 473
QY 134 LPLSAFTNPTC--EIVDEKTVVHTSOTPV 162
DB 474 ----ALKNFSLGESSIGLTSVLSGSPFD 500

RESULT 14

US-08-893-333-2
Sequence 2, Application US/08893333A
Patent No. 5981705

GENERAL INFORMATION:
APPLICANT: Kornbluth, Jacki
TITLE OF INVENTION: DNA Encoding Natural Killer Lytic Associated Protein
FILE REFERENCE: D5705CIP/D
CURRENT APPLICATION NUMBER: US/08/893,333A
CURRENT FILING DATE: 1997-07-16
NUMBER OF SEQ ID NOS: 17
SEQ ID NO 2
LENGTH: 587
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Amino acid sequence of Natural Killer Lytic
OTHER INFORMATION: Associated Protein encoded by nucleotides
OTHER INFORMATION: 190 to 1953 of Sequence ID. No. 5981705 1
Patent No. 5981705
US-08-893-333-2

Query Match 8.1%; Score 74.5; DB 2; Length 587;
Best Local Similarity 25.2%; Pred. No. 5;
Matches 38; Conservative 17; Mismatches 39; Indels 57; Gaps 8;
QY 52 ALPARSAAVQPIVIGISORV-----RMSKEKDLGTLGYVLGITMVYII----- 97
DB 367 ALPAM-----VIGIPVYGRKHSRYEGRKTSKHKRLALTG--GTTLSVIASPV 415
QY 98 ----IAIGAGIILGYSY-----KRGKDIK---EHDQKVCEREMQRT 133
DB 416 AAVSVGIGVPMILAYVGVVPISLCRGGYGVSTANGKVGKIERDEDDGPITVADMR-- 473
QY 134 LPLSAFTNPTC--EIVDEKTVVHTSQTPTVD 162
DB 474 ----ALKNPSIGESSIGLTSVLSTSGSPTD 500

RESULT 15

US-08-882-046-4
; Sequence 4, Application US/08882046
; Patent No. 6136952
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; APPLICANT: Hood, Leroy
; APPLICANT: Krantz, Ian D.
; APPLICANT: Spinner, Nancy B.
; TITLE OF INVENTION: Human Jagged Polypeptide, Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,046
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UW 2637
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1148 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-882-046-4

Query Match 8.1%; Score 74; DB 4; Length 1148;
Best Local Similarity 22.0%; Pred. No. 15;
Matches 54; Conservative 20; Mismatches 77; Indels 94; Gaps 11;

QY 2 GEAGV--PEKRPC-----EDLRC-----PETTS-----QALPAFTTEIQ 33
DB 858 GEGCAEPPSTPCIPRSGHLDNNCARLTHFNDRDHPQGTTVAGISGIRSLPATRAVAR 917
QY 34 E-----ASEGPGADEVOV--FAPANALPARS--EAAAVQPIVIGISORVRNMSKEKDD 81
DB 918 DRLLVLCDRASSGASAVEAVSFSPARDLPDSSLIOGAHAIVAAITDR----- 967
QY 82 LGTLGYVLGITMVIIAIGAGIILG-----YSYKRGDLKE 118

DB 968 -GNSLLLAIVIEVKETVVTGGSTGLLVPLCGAFSVLMACVVLQVMTFRKRK----- 1022
QY 119 QHDQKVCEREMQRT-----LPLSAFTNPTCEIYDEKTVVHTSQTPTVDPEQSGSTP 169
DB 1023 -----ERERERLPRRESSANNQMAPLNPINPIERPGHGDVLYQCCNFTPPRRRDAEA 1075
QY 170 IMGQA 174
DB 1076 LEPGA 1080

Search completed: April 7, 2003, 09:22:33
Job time : 20.8628 secs

